

Human albumin (HA)
Human albumin (HA)
Human serum albumin
Human serum albumin
B Lymphocyte stimu
Human B Lymphocyte
Mature form of hum
Recombinant human
Pre human serum al
Chimeric human ser
Recombinant human
Peptide-HSA-G-CSF c
G-CSF-(GLY)4-HSA c
HSA-WVF(417-413) c
Sequence of mature
Mature protein of
Human serum albumin
Sequence of human
Human serum albumin
Cancer metastasis
Human serum albumin
Human serum albumin
Human albumin. Ho
Myosin light chain
Human serum albumin
Human serum albumin
Human serum albumin
Human serum albumin
Sequence of prepro
Human serum albumin
HSA. Picilia pasto
HSA:Fc gamma RIIF f
Human serum albumin

PS Disclosure; fig 2; 20pp; English.

CC Mature protein of human serum albumin (see corresp. AAN90128).
CC Used to make new N-terminal fragments which are used as plasma
CC expanders, or as substitutes for HSA or BSA, in tissue culture
CC media.

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 DAHSEVAHRPKDGEENFKALVLIARQYIQCCPFEDHVKLVNEVEFAKTCVADESAAE 60
DB 1 DAHSEVAHRPKDGEENFKALVLIARQYIQCCPFEDHVKLVNEVEFAKTCVADESAAE 60
OY 61 NCDSLSLHTLFGDKICTVATLTRETYGEMADCCAKOPEPNECFLOHKDNPMLPRLVREY 120
DB 61 NCDSLSLHTLFGDKICTVATLTRETYGEMADCCAKOPEPNECFLOHKDNPMLPRLVREY 120
OY 121 DVMTAFHDNEETFLKLYEIAARRHPYFVAPPELLFFAKRYKAAFTCCQAADRAACLLP 180
DB 121 DVMTAFHDNEETFLKLYEIAARRHPYFVAPPELLFFAKRYKAAFTCCQAADRAACLLP 180
OY 121 DVMTAFHDNEETFLKLYEIAARRHPYFVAPPELLFFAKRYKAAFTCCQAADRAACLLP 180
DB 121 DVMTAFHDNEETFLKLYEIAARRHPYFVAPPELLFFAKRYKAAFTCCQAADRAACLLP 180
OY 181 KLDELROGKASSAKQRLKASLOKFGERRAKAMAVARLSQRPFAEFAEYSKLVTDLTK 240
DB 181 KLDELROGKASSAKQRLKASLOKFGERRAKAMAVARLSQRPFAEFAEYSKLVTDLTK 240
OY 241 VHTCCGHDLLECDRDLAKYICENODSISSKLECCERPLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLECDRDLAKYICENODSISSKLECCERPLEKSHCIAEVENDEMPA 300
OY 241 VHTCCGHDLLECDRDLAKYICENODSISSKLECCERPLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLECDRDLAKYICENODSISSKLECCERPLEKSHCIAEVENDEMPA 300
OY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMLFYEARHPDYSVLLRLAKTYETTLTKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMLFYEARHPDYSVLLRLAKTYETTLTKC 360
OY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMLFYEARHPDYSVLLRLAKTYETTLTKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMLFYEARHPDYSVLLRLAKTYETTLTKC 360
OY 361 CAADPHCEYAKVDEFPVLEEPONLIKONCELEFQEGEYKFNALLVRYTKKVPVST 420
DB 361 CAADPHCEYAKVDEFPVLEEPONLIKONCELEFQEGEYKFNALLVRYTKKVPVST 420
OY 421 PTLVEVSNTLKGVSCKCKHPBAKRMPCAEYLSVNLQCLVLEHKTVPVSRYTKCTES 480
DB 421 PTLVEVSNTLKGVSCKCKHPBAKRMPCAEYLSVNLQCLVLEHKTVPVSRYTKCTES 480
OY 481 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSKERQIKQIALVELVYKHKPKAT 540
DB 481 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSKERQIKQIALVELVYKHKPKAT 540
OY 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEGSKLVAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEGSKLVAASQAALGL 585

```

RESULT 2

AAR05318

ID AAR05318 standard; protein; 585 AA.

AC AAR05318;

DT 08-OCT-1990 (first entry)

XX Human serum albumin gene product.

XX Human serum albumin; HSA-A; yeast; ds.

XX Homo sapiens.

XX JP02117384-A.

XX 01-MAY-1990.

XX 26-OCT-1988; 88JP-0268302.

XX 26-OCT-1988; 88JP-0268302.

XX (TOFU) TOA NENRYO KOGYO KK.

XX WPI: 1990-176228/23.

XX N-PSDB; AAO04719.

XX Human serum albumin prepn. by yeast host -
XX by culturing transformed plasmid yeast to produce serum, and
XX removing it.

XX Disclosure; ; pp; Japanese.

CC Mature HSA-A may be produced using the sequence incorporated into a
CC plasmid vector with suitable controllers, and transferred to a yeast
CC expression system.

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 DAHSEVAHRPKDGEENFKALVLIARQYIQCCPFEDHVKLVNEVEFAKTCVADESAAE 60
DB 1 DAHSEVAHRPKDGEENFKALVLIARQYIQCCPFEDHVKLVNEVEFAKTCVADESAAE 60
OY 61 NCDSLSLHTLFGDKICTVATLTRETYGEMADCCAKOPEPNECFLOHKDNPMLPRLVREY 120
DB 61 NCDSLSLHTLFGDKICTVATLTRETYGEMADCCAKOPEPNECFLOHKDNPMLPRLVREY 120
OY 121 DVMTAFHDNEETFLKLYEIAARRHPYFVAPPELLFFAKRYKAAFTCCQAADRAACLLP 180
DB 121 DVMTAFHDNEETFLKLYEIAARRHPYFVAPPELLFFAKRYKAAFTCCQAADRAACLLP 180
OY 121 DVMTAFHDNEETFLKLYEIAARRHPYFVAPPELLFFAKRYKAAFTCCQAADRAACLLP 180
DB 121 DVMTAFHDNEETFLKLYEIAARRHPYFVAPPELLFFAKRYKAAFTCCQAADRAACLLP 180
OY 181 KLDELROGKASSAKQRLKASLOKFGERRAKAMAVARLSQRPFAEFAEYSKLVTDLTK 240
DB 181 KLDELROGKASSAKQRLKASLOKFGERRAKAMAVARLSQRPFAEFAEYSKLVTDLTK 240
OY 241 VHTCCGHDLLECDRDLAKYICENODSISSKLECCERPLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLECDRDLAKYICENODSISSKLECCERPLEKSHCIAEVENDEMPA 300
OY 241 VHTCCGHDLLECDRDLAKYICENODSISSKLECCERPLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLECDRDLAKYICENODSISSKLECCERPLEKSHCIAEVENDEMPA 300
OY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMLFYEARHPDYSVLLRLAKTYETTLTKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMLFYEARHPDYSVLLRLAKTYETTLTKC 360
OY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMLFYEARHPDYSVLLRLAKTYETTLTKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMLFYEARHPDYSVLLRLAKTYETTLTKC 360
OY 361 CAADPHCEYAKVDEFPVLEEPONLIKONCELEFQEGEYKFNALLVRYTKKVPVST 420
DB 361 CAADPHCEYAKVDEFPVLEEPONLIKONCELEFQEGEYKFNALLVRYTKKVPVST 420
OY 421 PTLVEVSNTLKGVSCKCKHPBAKRMPCAEYLSVNLQCLVLEHKTVPVSRYTKCTES 480
DB 421 PTLVEVSNTLKGVSCKCKHPBAKRMPCAEYLSVNLQCLVLEHKTVPVSRYTKCTES 480
OY 481 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSKERQIKQIALVELVYKHKPKAT 540
DB 481 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSKERQIKQIALVELVYKHKPKAT 540
OY 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEGSKLVAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDEKTCFAEGSKLVAASQAALGL 585

```

RESULT 3

AAR08457

AC AAR08457;

DT 16-APR-1991 (first entry)

```

DE Human serum albumin.
XX HSA; folding; ss.
XX Homo sapiens.
XX Key
XX Region 123..303
XX Region 1..303
XX Region /label= A
XX Region /label= B
XX Region 123..585
XX Region /label= C
XX JP02227079-A.
XX 25-AUG-1989.
XX 10-SEP-1990: 90JP-0250926.
XX 06-OCT-1988: 88JP-0250926.
XX (TOFU ) TONEN CORP.
XX WPI: 1990-317325/42.
XX N-PSDB: AAQ06099.
XX New human serum albumin fragments - used to bond medicines and for
XX stable folding of protein(s).
XX Claim 1; Fig 8; 24pp; Japanese.
XX Fragments A-C of HSA are expressed as fusion proteins with the
XX signal peptide of E. coli alkaline phosphatase. The fragments are
XX selected for their specific properties. The C-terminal truncated
XX fragment, B, does not bind long-chain fatty acids but does bind to
XX various medicines at the central region. The N-terminal truncated
XX fragment, C, has good stability in protein folding. The central
XX segment, A, has characteristics of both B and C.
XX See also AAQ06096-006098.
XX
XX Sequence 585 AA:
XX
Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKEVAHFRKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEYTEFAKTCVADESAAE 60
DB 1 DAHKEVAHFRKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEYTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVAPRY 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVAPRY 120
QY 121 DWACSTAFHNEFTFKKLYETARHHPYAPPELLEFAKRYKAFFECOAADKACCLIP 180
DB 121 DWACSTAFHNEFTFKKLYETARHHPYAPPELLEFAKRYKAFFECOAADKACCLIP 180
QY 121 DWACSTAFHNEFTFKKLYETARHHPYAPPELLEFAKRYKAFFECOAADKACCLIP 180
DB 121 DWACSTAFHNEFTFKKLYETARHHPYAPPELLEFAKRYKAFFECOAADKACCLIP 180
QY 181 KLDELDEKASSAKORLKCASLOKGERAFKAMAVARLSQRPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEKASSAKORLKCASLOKGERAFKAMAVARLSQRPKAEFAEVSKLVTDLTK 240
QY 241 VITECHGDLLECADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAEYENDMPA 300
DB 241 VITECHGDLLECADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAEYENDMPA 300
QY 301 DLPSLAADVESKDVCKNTAAEKDVFLEGMFLYEYARRHDPYSVLLRLAKYETLEK 360
DB 301 DLPSLAADVESKDVCKNTAAEKDVFLEGMFLYEYARRHDPYSVLLRLAKYETLEK 360
QY 361 CAADHDECYAKVDEFFKPLVEEPQNLIKONCELFQOLGEYKRONALLVRYTKVQVST 420
DB 361 CAADHDECYAKVDEFFKPLVEEPQNLIKONCELFQOLGEYKRONALLVRYTKVQVST 420

```

```

DB 361 CAADHDECYAKVDEFFKPLVEEPQNLIKONCELFQOLGEYKRONALLVRYTKVQVST 420
QY 421 PTLVEYSRNIGKSGKCKHPEAKRMPACADYLSVVLNOLCVLHETTPVSDRTKCTES 480
DB 421 PTLVEYSRNIGKSGKCKHPEAKRMPACADYLSVVLNOLCVLHETTPVSDRTKCTES 480
QY 481 LVNRRPCFSALDEVDYVPEKFEPAETFTFHADICTSEKEROIKOTALVELVHKRPAT 540
DB 481 LVNRRPCFSALDEVDYVPEKFEPAETFTFHADICTSEKEROIKOTALVELVHKRPAT 540
QY 541 KEOLKAVMDFAAFVEKCKKADDKETCFABEGKRLVAASQALGL 585
DB 541 KEOLKAVMDFAAFVEKCKKADDKETCFABEGKRLVAASQALGL 585

RESULT 4
AAR80301
ID AAR80301 standard; Protein; 585 AA.
XX
AC AAR80301;
XX
DT 17-JAN-1996 (first entry)
XX
DE Human serum albumin.
XX
KW Serum albumin; HSA; aspartyl protease-3; Yap3p;
XX Saccharomyces cerevisiae.
XX
OS Homo sapiens.
XX
PN W09523857-A1.
XX
PD 08-SEP-1995.
XX
PF 01-MAR-1995; 95WO-GB00434.
XX
PR 05-MAR-1994; 94GB-0004270.
XX
PA (DELTA ) DELTA BIOTECHNOLOGY LTD.
XX
XX Gilbert SC, Kerry-Williams SM;
XX
DR WPI: 1995-320572/41.
XX
N-PSDB: AAQ98695.
XX
PT Yeast with reduced levels of aspartyl protease 3 proteolytic
XX activity - used to secrete human albumin without prodn. of the 45
XX kd fragment
XX
PS Example 1; Page 26-28; 50pp; English.
XX
XX The cDNA given in AAQ98695, which encodes HSA (AAR80301), was subjected
XX to site-directed mutagenesis to investigate the role of
XX endoproteases in the generation of a 45 kDa albumin fragment obtd.
XX when the cDNA is expressed in S. cerevisiae. Mutations were: R410A;
XX I407A, I408Y, V409A; and R410A, K413Q, K414Q. The latter set of
XX mutations, especially, improved stability of HSA to yeast Yap3p
XX proteolytic cleavage, allowing increased prodn. of recombinant HSA.
XX
XX Sequence 585 AA:
XX
Query Match 100.0%; Score 3103; DB 16; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKEVAHFRKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEYTEFAKTCVADESAAE 60
DB 1 DAHKEVAHFRKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEYTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVAPRY 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLVAPRY 120

```

QY 121 DVMCTAFHDNEETFLKKYLYEIAIRRHPIFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
 Db 121 DVMCTAFHDNEETFLKKYLYEIAIRRHPIFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
 QY 181 KLDELREGKASSAKORLKCSLQKFGERRAKAANAVALSORFPKAEFAEVSKLVTDLTK 240
 Db 181 KLDELREGKASSAKORLKCSLQKFGERRAKAANAVALSORFPKAEFAEVSKLVTDLTK 240
 QY 241 VHTCCGHDLEECADDDRADLAKYICENODSISSKLKECCEKPLEKSHCIAVENDEMPA 300
 Db 241 VHTCCGHDLEECADDDRADLAKYICENODSISSKLKECCEKPLEKSHCIAVENDEMPA 300
 QY 301 DLPSLADFEVSKDYCKNYAKADVFLGMLFYEARRRHDYSVLLRLAKTYETTLK 360
 Db 301 DLPSLADFEVSKDYCKNYAKADVFLGMLFYEARRRHDYSVLLRLAKTYETTLK 360
 QY 361 CAADPHCEYAKVDEKRPVLEEPONLIKONCELEBOLGEYKFNALLVRYTKVPVOST 420
 Db 361 CAADPHCEYAKVDEKRPVLEEPONLIKONCELEBOLGEYKFNALLVRYTKVPVOST 420
 QY 421 PTLVEFSRNIGKVGSKCKHPEAKRMPCAEEDYLSVNLQCVLHEKTPVSDRYTKCCTES 480
 Db 421 PTLVEFSRNIGKVGSKCKHPEAKRMPCAEEDYLSVNLQCVLHEKTPVSDRYTKCCTES 480
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKERQIKQTALVELVHKPKAT 540
 Db 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKERQIKQTALVELVHKPKAT 540
 QY 541 KEOLKAVMDPFAAFVEKCKKADDEKTCFAEGSKTLVAASQAALGL 585
 Db 541 KEOLKAVMDPFAAFVEKCKKADDEKTCFAEGSKTLVAASQAALGL 585
 RESULT 5
 AAO20111
 ID AAO20111 standard; Protein: 585 AA.
 AC AAO20111;
 DT 06-AUG-2002 (first entry)
 DE HSA protein sequence related to the growth hormone protein.
 KW Serum albumin-growth hormone fusion protein; growth hormone;
 KM Down's syndrome.
 XX
 OS Unidentified.
 PN KR9076789-A.
 PD 15-OCT-1999.
 PF 25-JUN-1998; 98KR-0704914.
 PR 30-DEC-1995; 95GB-0026733.
 PR 19-DEC-1996; 96WO-GB03164.
 PA (DELZ) DELTA BIOTECHNOLOGY LTD.
 XX WPI: 1997-363680/55.
 DR N-PSDB; AAK99568.
 XX
 PT Serum albumin-growth hormone fusion protein - useful to treat growth
 PT hormone related diseases, e.g. Down's syndrome
 XX
 PS Disclosure: Fig 6; 21pp: Korean.
 CC The invention relates to a serum albumin-growth hormone fusion protein -
 CC useful to treat growth hormone related diseases such as Down's syndrome.
 CC This sequence represents a HSA protein related to the serum albumin-
 CC growth hormone protein of the invention.
 XX
 SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 18; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHRKFDLGENFRALVILAFAYOIQOCPFEDHVLVNEVEFAKTCVADSAAE 60
 Db 1 DAHSEVAHRKFDLGENFRALVILAFAYOIQOCPFEDHVLVNEVEFAKTCVADSAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKPERNECEFLQHKDNPMLPLVPREV 120
 Db 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKPERNECEFLQHKDNPMLPLVPREV 120
 QY 121 DVMCTAFHDNEETFLKKYLYEIAIRRHPIFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
 Db 121 DVMCTAFHDNEETFLKKYLYEIAIRRHPIFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
 QY 181 KLDELREGKASSAKORLKCSLQKFGERRAKAANAVALSORFPKAEFAEVSKLVTDLTK 240
 Db 181 KLDELREGKASSAKORLKCSLQKFGERRAKAANAVALSORFPKAEFAEVSKLVTDLTK 240
 QY 241 VHTCCGHDLEECADDDRADLAKYICENODSISSKLKECCEKPLEKSHCIAVENDEMPA 300
 Db 241 VHTCCGHDLEECADDDRADLAKYICENODSISSKLKECCEKPLEKSHCIAVENDEMPA 300
 QY 301 DLPSLADFEVSKDYCKNYAKADVFLGMLFYEARRRHDYSVLLRLAKTYETTLK 360
 Db 301 DLPSLADFEVSKDYCKNYAKADVFLGMLFYEARRRHDYSVLLRLAKTYETTLK 360
 QY 361 CAADPHCEYAKVDEKRPVLEEPONLIKONCELEBOLGEYKFNALLVRYTKVPVOST 420
 Db 361 CAADPHCEYAKVDEKRPVLEEPONLIKONCELEBOLGEYKFNALLVRYTKVPVOST 420
 QY 421 PTLVEFSRNIGKVGSKCKHPEAKRMPCAEEDYLSVNLQCVLHEKTPVSDRYTKCCTES 480
 Db 421 PTLVEFSRNIGKVGSKCKHPEAKRMPCAEEDYLSVNLQCVLHEKTPVSDRYTKCCTES 480
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKERQIKQTALVELVHKPKAT 540
 Db 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKERQIKQTALVELVHKPKAT 540
 QY 541 KEOLKAVMDPFAAFVEKCKKADDEKTCFAEGSKTLVAASQAALGL 585
 Db 541 KEOLKAVMDPFAAFVEKCKKADDEKTCFAEGSKTLVAASQAALGL 585
 RESULT 6
 AAY84873
 ID AAY84873 standard; Protein: 585 AA.
 AC AAY84873;
 DT 08-AUG-2000 (first entry)
 DE Amino acid sequence of a human albumin protein.
 KW Human; albumin; ischemic state; serum protein; metal ion salt;
 KW perioperative ischemia; ischemia; myocardial infarction;
 KW progressive coronary artery disease.
 XX
 OS Homo sapiens.
 PN
 PD
 PF
 PT Key Location/Qualifiers
 PT Modified-site 1 /note="optionally acetylated, and claimed under
 FT claim 56"
 FT
 XX WO200020840-A1.
 XX 13-APR-2000.
 XX 01-OCT-1999; 99WO-US22905.


```

PR 02-OCT-1998; 98US-0102738.
PR 02-OCT-1998; 98US-0165581.
PR 02-OCT-1998; 98US-0165926.
PR 11-JAN-1999; 99US-0115392.
XX
PA (ISCH-) ISCHEMIA TECHNOLOGIES INC.
XX
PI Bar-Or D, Lau E, Winkler JV;
XX
XX WPI: 2000-303843/26.
XX
XX New method for the continuous detection of ischemic states comprises
PT detecting and quantifying the existence of an alteration of the serum
PT protein albumin -
XX
XX Disclosure; Page 97-100; 105pp; English.
XX
XX The present sequence represents human albumin protein. The specification
CC describes a method for the continuous detection of ischemic states. The
CC method comprises detecting and quantifying the existence of an alteration
CC of the serum protein albumin. The method comprises contacting a
CC biological sample containing albumin from the patient with an excess
CC quantity of a metal ion salt, where the metal ion binds to the N-terminus
CC of naturally occurring human albumin, to form a mixture containing bound
CC metal ions and unbound metal ions, and then determining the amount of
CC metal ions bound to the albumin N-terminus. The amount of bound metal
CC ions is correlated to a known value to determine the occurrence or
CC non-occurrence of an ischemic event. The methods are useful for detection
CC of ischemic states. The methods are also useful for distinguishing
CC peroperative ischemia from ischemia caused by, amongst other things,
CC myocardial infarctions and progressive coronary artery disease.
XX
XX
SQ Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 21; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHHRFKDGEENFKALVLIIFAQYLQCCPFEDHVKLVNEVEFAKTCVADSAAE 60
DB 1 DAHSEVAHHRFKDGEENFKALVLIIFAQYLQCCPFEDHVKLVNEVEFAKTCVADSAAE 60
QY 61 NCDKSLHTLFGDKLCTVAATLRETYGEMADCCAKOEPERNBCFLOHKDNDNPLPRLVREY 120
DB 61 NCDKSLHTLFGDKLCTVAATLRETYGEMADCCAKOEPERNBCFLOHKDNDNPLPRLVREY 120
QY 121 DVMCSTAHNDNEFTFLKYLLEIARRHRYFYAPPELLFFAKRYKKAFTCCCAADKAACTLP 180
DB 121 DVMCSTAHNDNEFTFLKYLLEIARRHRYFYAPPELLFFAKRYKKAFTCCCAADKAACTLP 180
QY 181 KLDELRLREGKASSAKOKLKCAASLOKFGERAFAKMAVAARLSORPFAFAEVSKLVTDLTK 240
DB 181 KLDELRLREGKASSAKOKLKCAASLOKFGERAFAKMAVAARLSORPFAFAEVSKLVTDLTK 240
QY 241 VHTSECHGDLLECCADRADLAKYICENQDISISKLKCECKPRLLEKSHCIAEYENDMPA 300
DB 241 VHTSECHGDLLECCADRADLAKYICENQDISISKLKCECKPRLLEKSHCIAEYENDMPA 300
QY 301 DLPSLAADFYESKDYCKNVAEAKDVLGMFLYEYARRHPDY SVVLLRLAKTYETTLK 360
DB 301 DLPSLAADFYESKDYCKNVAEAKDVLGMFLYEYARRHPDY SVVLLRLAKTYETTLK 360
QY 361 CAADAPHEGAKYFDEKPLVEEPONLIKONCELFEDLGKFKFONALLVYTKVPOVST 420
DB 361 CAADAPHEGAKYFDEKPLVEEPONLIKONCELFEDLGKFKFONALLVYTKVPOVST 420
QY 421 PTLVEVSRLNGKYSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEVSRLNGKYSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPFSALVEDEYVPEFNAETTFHADICTLSEKERQIKKOTALVELVYHKRPAT 540
DB 481 LVNRRPFSALVEDEYVPEFNAETTFHADICTLSEKERQIKKOTALVELVYHKRPAT 540

```

```

QY 541 KEOLKAVMDPFAAFVEKCKKADKCTCFAEGRKLVAAASOALGL 585
DB 541 KEOLKAVMDPFAAFVEKCKKADKCTCFAEGRKLVAAASOALGL 585
XX
XX RESULT 7
XX ID AAY83946
XX AC AAY83946;
XX DT 28-JUL-2000 (first entry)
XX
XX Yeast codon-biased recombinant human serum albumin protein.
XX
XX Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
XX overlapping oligonucleotide; expression vector.
XX
XX Homo sapiens.
XX Synthetic.
XX CN1239103-A.
XX
XX 22-DEC-1999.
XX
XX 17-JUN-1998; 98CN-0102506.
XX
XX 17-JUN-1998; 98CN-0102506.
XX
XX 17-JUN-1998; 98CN-0102506.
XX
XX (HAUJ-) HAUJ BIOENGINEERING CO LTD.
XX
XX Li S, Lu D;
XX
XX WPI: 2000-351198/31.
XX
XX N-PSDB; AAA10091.
XX
XX Process for preparing recombinant human serum albumin - which comprises
PT yeast biased sex codons
PT
XX Disclosure; Fig 1; 44p; Chinese.
XX
XX The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to
CC comprise a yeast codon bias. The complete HSA gene (AAA10091) was
CC generated as three synthetic fragments (AAA10092-10094) joined by
CC recombinant DNA technology. Each HSA fragment was synthesised from
CC overlapping oligonucleotide fragments that were extended. This sequence
CC represents the complete sequence of the HSA encoded by the human gene
CC with a yeast codon bias. The invention also covers a recombinant
CC expression vector, yeast host cells carrying the recombinant expression
CC vector and the process for producing human serum albumin in the yeast
CC host cell, especially in secretory mode.
XX
XX
SQ Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 21; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHHRFKDGEENFKALVLIIFAQYLQCCPFEDHVKLVNEVEFAKTCVADSAAE 60
DB 1 DAHSEVAHHRFKDGEENFKALVLIIFAQYLQCCPFEDHVKLVNEVEFAKTCVADSAAE 60
QY 61 NCDKSLHTLFGDKLCTVAATLRETYGEMADCCAKOEPERNBCFLOHKDNDNPLPRLVREY 120
DB 61 NCDKSLHTLFGDKLCTVAATLRETYGEMADCCAKOEPERNBCFLOHKDNDNPLPRLVREY 120
QY 121 DVMCSTAHNDNEFTFLKYLLEIARRHRYFYAPPELLFFAKRYKKAFTCCCAADKAACTLP 180
DB 121 DVMCSTAHNDNEFTFLKYLLEIARRHRYFYAPPELLFFAKRYKKAFTCCCAADKAACTLP 180
QY 181 KLDELRLREGKASSAKOKLKCAASLOKFGERAFAKMAVAARLSORPFAFAEVSKLVTDLTK 240

```

```
Db      181 KILDELREGKASAKORLKASLOKFGERRAKAMAVARLSQRPFAEFAVSKLVTDLTK 240
Qy      241 VHTCCGDDLECCADDDRADLAKYICENDOSTISSKLECCCEPPLKESKICAEVENDEMPA 300
Db      241 VHTCCGDDLECCADDDRADLAKYICENDOSTISSKLECCCEPPLKESKICAEVENDEMPA 300
Qy      301 DLPSIADPFVESKDVCKNVAEAKDVFLGMLFYEARRRHDYSVYLLRLAKTYETLEKC 360
Db      301 DLPSIADPFVESKDVCKNVAEAKDVFLGMLFYEARRRHDYSVYLLRLAKTYETLEKC 360
Qy      361 CAADPFHECYAKVDERKPLVEEPONLIKONCEPEQJGGEKFGONALLVRTKKVPOVST 420
Db      361 CAADPFHECYAKVDERKPLVEEPONLIKONCEPEQJGGEKFGONALLVRTKKVPOVST 420
Qy      421 PTIVEVSNNLKVSKCKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPDSRVTKCTES 480
Db      421 PTIVEVSNNLKVSKCKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPDSRVTKCTES 480
Qy      481 LVNRRPCSALEVEDETVYVPKFENMETFTFHADICTLSEKEROIKKOTALVELVKKHPRAT 540
Db      481 LVNRRPCSALEVEDETVYVPKFENMETFTFHADICTLSEKEROIKKOTALVELVKKHPRAT 540
Qy      541 KEQLKAVDDPAAPEKCCCKADDKETGFAEGSKYVAASQAALGL 585
Db      541 KEQLKAVDDPAAPEKCCCKADDKETGFAEGSKYVAASQAALGL 585

RESULT 8
ID      ABB79006
XX      ABB79006 standard; Protein: 585 AA.
AC      ABB79006;
XX      ABB79006;
DT      01-AUG-2002 (first entry)
XX      01-AUG-2002 (first entry)
DE      Human mature albumin protein SEQ ID NO:18.
XX      Human; growth hormone; hGH; albumin; human serum albumin; HSA;
KW      albumin fusion protein; cytosstatic; anorectic; immunosuppressive;
KW      antidiabetic; antirheumatic; antiarthritic; psoriatic; cancer;
KW      non-Hodgkin's lymphoma; obesity; transplant rejection; psoriasis;
KW      type 1 diabetes mellitus; rheumatoid arthritis.
XX      Homo sapiens.
OS      Homo sapiens.
FH      Key
FT      Domain
FT      Location/Qualifiers
FT      1..194
FT      /label= 1
FT      Domain
FT      1..105
FT      /label= subdomain
FT      Disulfide-bond
FT      53..62
FT      Disulfide-bond
FT      75..91
FT      Disulfide-bond
FT      90..101
FT      Disulfide-bond
FT      106..119
FT      /note= "flexible inter-subdomain linker region"
FT      Domain
FT      120..194
FT      /label= subdomain
FT      Disulfide-bond
FT      124..169
FT      Disulfide-bond
FT      168..177
FT      Domain
FT      195..387
FT      /label= 2
FT      Domain
FT      195..291
FT      /label= subdomain
FT      Disulfide-bond
FT      245..253
FT      Disulfide-bond
FT      265..279
FT      Disulfide-bond
FT      278..289
FT      Disulfide-bond
FT      292..315
FT      /note= "flexible inter-subdomain linker region"
FT      Domain
FT      316..387
FT      /label= subdomain
FT      Disulfide-bond
FT      316..361
FT      Disulfide-bond
FT      360..369
```

```
FT      Domain
FT      388..585
FT      /label= 3
FT      Domain
FT      388..491
FT      /label= subdomain
FT      Disulfide-bond
FT      392..438
FT      Disulfide-bond
FT      437..448
FT      Disulfide-bond
FT      461..477
FT      Disulfide-bond
FT      476..487
FT      Disulfide-bond
FT      492..511
FT      /note= "flexible inter-subdomain linker region"
FT      Domain
FT      512..585
FT      /label= subdomain
FT      Disulfide-bond
FT      514..559
FT      Disulfide-bond
FT      558..567
FT      WO200179442-A2.
FT      25-OCT-2001.
FT      12-APR-2001; 2001WO-US11850.
FT      12-APR-2000; 2000US-229358P.
FT      25-APR-2000; 2000US-199384P.
FT      21-DEC-2000; 2000US-256931P.
FT      (HUMA-) HUMAN GENOME SCT INC.
FT      PA
FT      Rosen CA, Haseltine WA.
FT      WPI: 2001-611723/70.
FT      DR      N-PSDB; ABB67288.
FT      PT      New albumin fusion proteins, useful for treating diseases and disorders
FT      such as cancer, comprise therapeutic protein fused to albumin
FT      PS      Claim 1; Fig 11; 413pp; English.
XX      The present invention describes an albumin fusion protein (1) comprising
XX      a therapeutic protein: X and (a fragment or variant of) albumin
XX      comprising a the fully defined sequence in ABB79006 of 585 amino acids,
XX      (where the fragment or variant has albumin or therapeutic protein: X
XX      activity). (1) can have cytosstatic, anorectic, immunosuppressive,
XX      antidiabetic, antirheumatic, antiarthritic and psoriatic activities.
XX      Albumin fusion proteins are stabilised therapeutic proteins e.g
XX      antibodies to C5, C242 and CD80 useful for treating various diseases
XX      and disorders such as non-Hodgkin's lymphoma, cancer, obesity,
XX      transplant rejection, type 1 diabetes mellitus, rheumatoid arthritis
XX      and psoriasis. Fusing albumin to therapeutic proteins stabilises the
XX      therapeutic protein, extends the shelf life and retains the in vitro or
XX      in vivo biological activity. It also reduces the need to formulate
XX      protein solutions with large excesses of carrier proteins to prevent
XX      loss of therapeutic proteins due to factors such as binding to the
XX      container. The fusion proteins are easily dispensed with a simple
XX      formulation requiring minimal post storage manipulation. The fusion of
XX      therapeutic proteins to albumin confers stability in aqueous or other
XX      solution. The present sequence represents the mature human albumin (HA)
XX      protein which is used in the exemplification of the present invention.
XX      Sequence 585 AA:
XX      Query Match 100.0%; Score 3103; DB 22; Length 585;
XX      Best Local Similarity 100.0%; Pred. No. 1e-254;
XX      Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 DAHKESEVAFKFDLSEENFKALVLLAFAYOYLQCCPFEDHVKLVNEVTERAKTCADESAE 60
Db      1 DAHKESEVAFKFDLSEENFKALVLLAFAYOYLQCCPFEDHVKLVNEVTERAKTCADESAE 60
Qy      61 NCCKSLHTLFGDKICTVATLTRETYGBMADCCAKQBERNECFLQKNDPNPLPRLYRREV 120
Db      61 NCCKSLHTLFGDKICTVATLTRETYGBMADCCAKQBERNECFLQKNDPNPLPRLYRREV 120
Qy      121 DVMCTAFHNDNETPLKKYLYETIARRHPYFYAELLFFPAKRYKAAPTECCQAADKAACLLP 180
Db      121 DVMCTAFHNDNETPLKKYLYETIARRHPYFYAELLFFPAKRYKAAPTECCQAADKAACLLP 180
```

```

Db      121 DVMCTAFHNDNEETFLKKYLYETIARRHPYAPPELLFPARKRYKAFTCCOADKAACLLP 180
QY      181 KIDELRDEGKASSAKRLKASLQKFGGERAFKAMAVARLSQRPKAEFAVSKLYVDLTK 240
Db      181 KIDELRDEGKASSAKRLKASLQKFGGERAFKAMAVARLSQRPKAEFAVSKLYVDLTK 240
QY      241 VTECHGDLLECCADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAEYENDMPA 300
Db      241 VTECHGDLLECCADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAEYENDMPA 300
QY      301 DEPSLADEVESKDVCKNYAEAKDVELGMEFLYETARRHPDYSVLLRLAKYETTLK 360
Db      301 DEPSLADEVESKDVCKNYAEAKDVELGMEFLYETARRHPDYSVLLRLAKYETTLK 360
QY      361 CAADHDEGKAYVDFEFKPLVEBPQNLKONCELFQOLGKYFQNALVRYTKKPVOST 420
Db      361 CAADHDEGKAYVDFEFKPLVEBPQNLKONCELFQOLGKYFQNALVRYTKKPVOST 420
QY      421 PTLVEYSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCCTES 480
Db      421 PTLVEYSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCCTES 480
QY      481 LVNRRPCFSALEVEDETVYRKFNATFTFHADICTLSEKERQIKKOTALVELYKHKPKAT 540
Db      481 LVNRRPCFSALEVEDETVYRKFNATFTFHADICTLSEKERQIKKOTALVELYKHKPKAT 540
QY      541 KQOLKAVMDFAFVYKCKCKADDKETCFABEGKKLYAASQALGL 585
Db      541 KQOLKAVMDFAFVYKCKCKADDKETCFABEGKKLYAASQALGL 585

```

RESULT 9

AAE13399 standard; Protein; 585 AA.

```

ID      AAE13399 standard; Protein; 585 AA.
XX      AAE13399;
XX      12-FEB-2002 (first entry)
DT      Human albumin (HA) protein.
XX      Human: albumin; HA: fusion protein; immune system disorder; syphilis;
XX      transplant rejection; blood related disorder; myocardial infarction;
XX      hyperproliferative disorder; acute myeloid leukemia; renal disorder;
XX      glomerulonephritis; cardiovascular disease; arrhythmia; rhinitis;
XX      respiratory disorder; rheocytocytoma; reproductive system disorder;
XX      measles; gastrointestinal disorder; irritable bowel syndrome; HIV;
XX      human immunodeficiency virus; wound healing; renal cell carcinoma;
XX      melanoma; gene therapy.
OS      Homo sapiens.
XX      Key
FH      Location/Qualifiers
FT      54..61
FT      /label= Loop_I
FT      76..89
FT      /label= Loop_II
FT      92..100
FT      /label= Loop_III
FT      170..176
FT      /label= Loop_IV
FT      247..252
FT      /label= Loop_V
FT      266..277
FT      /label= Loop_VI
FT      280..288
FT      /label= Loop_VII
FT      362..368
FT      /label= Loop_VIII
FT      439..447
FT      /label= Loop_IX
FT      Domain

```

```

FT      Domain
FT      461..475
FT      /label= Loop_X
FT      Domain
FT      478..486
FT      /label= Loop_XI
FT      Domain
FT      560..566
FT      /label= Loop_XII
PN      WO200179258-A1.
XX      25-OCT-2001.
PD      12-APR-2001; 2001WO-US12008.
PF      12-APR-2000; 2000US-129358P.
PR      25-APR-2000; 2000US-199384P.
PR      21-DEC-2000; 2000US-256931P.
XX      (HUMA-) HUMAN GENOME SCI INC.
PA      (PRIN-) PRINCIPAL PHARM CORP.
PI      Rosen CA, Sadeghi H, Prior CP, Turner AJ;
XX      WPI; 2001-602931/68.
DR      N-PSDB; AAD22287.
XX      Claim 1; Fig 9; 325pp; English.
XX      The invention relates to albumin fusion proteins comprising therapeutic
XX      protein and human albumin (HA). The albumin fusion proteins are useful
XX      in the treatment, prevention, diagnosis, and/or detection of diseases,
XX      disorders such as immune system disorders (transplant rejection); blood
XX      related disorders (myocardial infarction); hyperproliferative disorders
XX      (childhood acute myeloid leukemia); renal disorder (glomerulonephritis);
XX      cardiovascular disorders (arrhythmias); respiratory disorders
XX      (non-allergic rhinitis); neurological diseases (Alzheimer's disease);
XX      endocrine disorders (pheocytocytoma); reproductive system disorders
XX      (syphilis); infectious diseases (measles); gastrointestinal disorders
XX      (irritable bowel syndrome) and wound healing. The albumin fusion
XX      proteins are also used in the treatment of metastatic renal cell
XX      carcinoma, metastatic melanoma, malignant melanoma and HIV (human
XX      immunodeficiency virus) infection. Nucleic acid encoding albumin fusion
XX      albumin (HA) protein.
XX      Sequence 585 AA:

```

```

SQ      Sequence 585 AA:
Query Match 100.0%; Score 3103; DB 22; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 DAHSEVARRFKDGLGENFKAVLYLFAQYLOCCPREDHYKLVNTEYFAKTCVADESA 60
Db      1 DAHSEVARRFKDGLGENFKAVLYLFAQYLOCCPREDHYKLVNTEYFAKTCVADESA 60
QY      61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEBERNECFLOHDDNDNPLRLVRPEY 120
Db      61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEBERNECFLOHDDNDNPLRLVRPEY 120
QY      121 DVMCTAFHNDNEETFLKKYLYETIARRHPYAPPELLFPARKRYKAFTCCOADKAACLLP 180
Db      121 DVMCTAFHNDNEETFLKKYLYETIARRHPYAPPELLFPARKRYKAFTCCOADKAACLLP 180
QY      181 KIDELRDEGKASSAKRLKASLQKFGGERAFKAMAVARLSQRPKAEFAVSKLYVDLTK 240
Db      181 KIDELRDEGKASSAKRLKASLQKFGGERAFKAMAVARLSQRPKAEFAVSKLYVDLTK 240
QY      241 VTECHGDLLECCADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAEYENDMPA 300
Db      241 VTECHGDLLECCADRADLAKYICENODSISSKLEKCECKPLLEKSHCIAEYENDMPA 300

```


KW wound healing; antiinflammatory; immunosuppressive; neuroprotective;
 KW cardiant; cytostatic; antileukemic; antirheumatic; antimicrobial;
 KW renal disorder.
 OS Homo sapiens.
 XX Key
 FH Location/Qualifiers
 FT 54..61
 FT /label= Loop-I
 FT 76..89
 FT /label= Loop-II
 FT 92..100
 FT /label= Loop-III
 FT 170..176
 FT /label= Loop-IV
 FT 247..252
 FT /label= Loop-V
 FT 266..277
 FT /label= Loop-VI
 FT 280..288
 FT /label= Loop-VII
 FT 362..368
 FT /label= Loop-VIII
 FT 439..447
 FT /label= Loop-IX
 FT 461..475
 FT /label= Loop-X
 FT 478..486
 FT /label= Loop-XI
 FT 560..566
 FT /label= Loop-XII
 PN WO200179443-A2.
 XX 25-OCT-2001.
 PD 12-APR-2001; 2001WO-US11924.
 PF 12-APR-2000; 2000US-229358P.
 PR 25-APR-2000; 2000US-199384P.
 PR 21-DEC-2000; 2000US-256931P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Haseltine WA;
 XX WPI; 2001-616754/71.
 DR N-PSDB; AAD21638.
 XX
 PT Albumin fusion proteins comprising a therapeutic protein and albumin;
 PT useful in the treating immune system disorders (e.g. transplant
 PT rejection), blood related disorders (e.g. myocardial infarction) and
 PT hyperproliferative disorders -
 XX
 PS Claim 1; Fig 9; 380pp; English.
 CC The invention relates to albumin fusion proteins comprising therapeutic
 CC protein and human albumin (HA). Therapeutic protein fused to albumin
 CC have an extended shelf-life. The albumin fusion proteins are useful in
 CC the treatment, prevention, diagnosis and/or detection of diseases,
 CC disorders such as immune system disorders (e.g. transplant rejection),
 CC blood related disorders (e.g. myocardial infarction), hyperproliferative
 CC disorders (e.g. childhood acute myeloid leukaemia), renal disorders
 CC (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),
 CC respiratory disorders (e.g. non-allergic rhinitis), neurological
 CC diseases (e.g. Alzheimer's disease), endocrine disorders (e.g.
 CC pheochromocytoma), reproductive system disorders (e.g. syphilis),
 CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
 CC irritable bowel syndrome) and wound healing. Nucleic acids encoding
 CC albumin fusion protein is used in gene therapy. The present sequence
 CC is human albumin (HA) protein.
 XX
 XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 22; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFRKDLGSENFKALVLIAPQYLQCCPEDHYKLVNEVEFAKTCVADESA 60
 DB 1 DAHSEVAHRFRKDLGSENFKALVLIAPQYLQCCPEDHYKLVNEVEFAKTCVADESA 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEBERNCEFTLOHKDNDNLPRLVREPV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEBERNCEFTLOHKDNDNLPRLVREPV 120
 QY 121 DVNCTAFHDEEFTLKKYLYEIRRRHPYFAPBLLEFPKRYKKAFTCCCAADKACLLP 180
 DB 121 DVNCTAFHDEEFTLKKYLYEIRRRHPYFAPBLLEFPKRYKKAFTCCCAADKACLLP 180
 QY 181 KLDELNDEGKASSAKORLKCASTLOKGERAFKMAVARLSQRPKAFPAVSKLVYDITK 240
 DB 181 KLDELNDEGKASSAKORLKCASTLOKGERAFKMAVARLSQRPKAFPAVSKLVYDITK 240
 QY 241 VHTCECHGDLLECCADRADLAKYICENODSISKLAECCERPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCECHGDLLECCADRADLAKYICENODSISKLAECCERPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLADAFVESKDVCKNVAEAKDVLGMFLYETARRHPYSVLLRLAKYETTLK 360
 DB 301 DLPSLADAFVESKDVCKNVAEAKDVLGMFLYETARRHPYSVLLRLAKYETTLK 360
 QY 361 CAADPHECYAKYFDEKRPVEEPONIKONCELFEOJGSEYKFNALVYTKVPQVST 420
 DB 361 CAADPHECYAKYFDEKRPVEEPONIKONCELFEOJGSEYKFNALVYTKVPQVST 420
 QY 421 PTLVEVSRNLGVSKCKHPEAKRMCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
 DB 421 PTLVEVSRNLGVSKCKHPEAKRMCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
 QY 481 LVNRRPCEFSALVEDETVYPRFEFNAEFTTFADICTISEKROIKQATVETLKHK 540
 DB 481 LVNRRPCEFSALVEDETVYPRFEFNAEFTTFADICTISEKROIKQATVETLKHK 540
 QY 541 KEOLKAVMDPFAFVEKCKADKRETCFADEGKKLVASQAALGL 585
 DB 541 KEOLKAVMDPFAFVEKCKADKRETCFADEGKKLVASQAALGL 585

RESULT 12
 AAE12403
 ID AAE12403 standard; Protein; 585 AA.
 AC AAE12403;
 XX 18-DEC-2001 (first entry)
 DE Human albumin (HA).
 XX Human; albumin; HA; immune system disorder; transplant rejection;
 KW blood related disorder; myocardial infarction; glomerulonephritis;
 KW hyperproliferative disorder; childhood acute myeloid leukaemia;
 KW renal cell carcinoma; cardiovascular disorder; vulnary; melanoma;
 KW arrhythmia; respiratory disorder; non-allergic rhinitis; antileukemic;
 KW neurological disease; Alzheimer's disease; endocrine disorder; measles;
 KW pheochromocytoma; reproductive system disorder; neuroprotective; syphilis;
 KW infectious disease; gastrointestinal disorder; wound healing; noctropic;
 KW irritable bowel syndrome; HIV; human immunodeficiency virus infection;
 KW cytostatic; antiinflammatory; gene therapy; immunosuppressive; cardiant;
 KW antiarrhythmic; antirheumatic; renal disorder; antimicrobial.
 XX Homo sapiens.
 OS
 XX Key
 FH Location/Qualifiers
 FT 54..61
 FT /label= Loop-I

FT Domain 76..89
 FT /label- Loop_II
 FT Domain 92..100
 FT /label- Loop_III
 FT Domain 170..176
 FT /label- Loop_IV
 FT Domain 247..252
 FT /label- Loop_V
 FT Domain 266..277
 FT /label- Loop_VI
 FT Domain 280..288
 FT /label- Loop_VII
 FT Domain 362..368
 FT /label- Loop_VIII
 FT Domain 439..447
 FT /label- Loop_IX
 FT Domain 461..475
 FT /label- Loop_X
 FT Domain 478..486
 FT /label- Loop_XI
 FT Domain 560..566
 FT /label- Loop_XII
 XX WO200179480-A1.
 PN 25-OCT-2001.
 PD 12-APR-2001; 2001WO-US11991.
 PE 12-APR-2001; 2000US-229358P.
 PR 25-APR-2000; 2000US-199384P.
 PR 21-DEC-2000; 2000US-256931P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Haseltine WA:
 PI WPI: 2001-616756/71.
 DR N-PSDB; AAD20005..
 XX Albumin fusion proteins comprising a therapeutic protein and albumin,
 PT useful in the treating metastatic renal cell carcinoma, metastatic
 PT melanoma, malignant melanoma, renal cell carcinoma, HIV (human
 PT immunodeficiency virus) or infection -
 XX
 PS Claim 1; Fig 9; 394pp; English.
 XX The invention relates to human albumin (HA) fusion proteins and their
 CC corresponding nucleic acid sequences. Therapeutic proteins fused to
 CC albumin or its fragments have an extended shelf-life. The albumin
 CC fusion proteins are useful in the treatment, prevention, diagnosis,
 CC and/or detection of diseases, disorders such as immune system
 CC disorders (e.g. transplant rejection), blood related disorders (e.g.
 CC myocardial infarction), hyperproliferative disorders (e.g. childhood
 CC acute myeloid leukaemia, metastatic renal cell carcinoma, metastatic
 CC melanoma, malignant melanoma, renal cell carcinoma), renal disorders
 CC (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),
 CC respiratory disorders (e.g. non-allergic rhinitis), neurological
 CC diseases (e.g. Alzheimer's disease), endocrine disorders (e.g.
 CC pheochromocytoma), reproductive system disorders (e.g. syphilis),
 CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
 CC irritable bowel syndrome), HIV (human immunodeficiency virus) infection
 CC and wound healing. Nucleic acids encoding albumin fusion protein is
 CC used in gene therapy. The present sequence is human albumin.
 XX
 SQ Sequence 585 AA:
 Query Match 100.0%; Score 3103; DB 22; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 DAHKEVAHRRFDLGEENFKALVILAFQYLQCCFPEHVKLVNTEVFAKTCVADSAAE 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1 DAHKEVAHRRFDLGEENFKALVILAFQYLQCCFPEHVKLVNTEVFAKTCVADSAAE 60
 OY NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLQHKDNDNPLPLVAPREV 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLQHKDNDNPLPLVAPREV 120
 OY DVMTAFHNDNEETFLKLYLVEIARRHPYFYAPDELLFFAKRYKAAFTCCQAADKACLLP 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 DVMTAFHNDNEETFLKLYLVEIARRHPYFYAPDELLFFAKRYKAAFTCCQAADKACLLP 180
 OY DVMTAFHNDNEETFLKLYLVEIARRHPYFYAPDELLFFAKRYKAAFTCCQAADKACLLP 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 DVMTAFHNDNEETFLKLYLVEIARRHPYFYAPDELLFFAKRYKAAFTCCQAADKACLLP 180
 OY KLDELREGKASSAKORLKCSLQKFGERRAKAAVAVRLSORPKAEFAEYSKLVTDLT 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 KLDELREGKASSAKORLKCSLQKFGERRAKAAVAVRLSORPKAEFAEYSKLVTDLT 240
 OY VHTCECHGDLLECADRDADLAKYICENODSISLTKBCECKPLLEKSHCIAEVENDEMPA 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 VHTCECHGDLLECADRDADLAKYICENODSISLTKBCECKPLLEKSHCIAEVENDEMPA 300
 OY DLPSLADEFVESKDVCKNVAEAKDVFLGMFLYEXARRHPDYSVYLLRLAKTYETLEKC 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 DLPSLADEFVESKDVCKNVAEAKDVFLGMFLYEXARRHPDYSVYLLRLAKTYETLEKC 360
 OY DLPSLADEFVESKDVCKNVAEAKDVFLGMFLYEXARRHPDYSVYLLRLAKTYETLEKC 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 DLPSLADEFVESKDVCKNVAEAKDVFLGMFLYEXARRHPDYSVYLLRLAKTYETLEKC 360
 OY CAADPHECYAKVDERKPLVEEPONLIKONCELFEOUGEYKFFONALLVRYTKKVPQVST 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 CAADPHECYAKVDERKPLVEEPONLIKONCELFEOUGEYKFFONALLVRYTKKVPQVST 420
 OY PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTPTVSDRYTKCCTES 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTPTVSDRYTKCCTES 480
 OY LVNRRPCFSALVEDETYVPKFPNAETFTFNADITLSEKEQIKOTALVAVLHNKRYAT 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 481 LVNRRPCFSALVEDETYVPKFPNAETFTFNADITLSEKEQIKOTALVAVLHNKRYAT 540
 OY LVNRRPCFSALVEDETYVPKFPNAETFTFNADITLSEKEQIKOTALVAVLHNKRYAT 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 541 LVNRRPCFSALVEDETYVPKFPNAETFTFNADITLSEKEQIKOTALVAVLHNKRYAT 540
 OY KEOLKAVDDFAFAVEKCKKADDKETCFPAEKGKLVAAASQALGL 585
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 541 KEOLKAVDDFAFAVEKCKKADDKETCFPAEKGKLVAAASQALGL 585
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 13
 AAE08578
 ID AAE08578 standard; Protein: 585 AA.
 XX
 AC AAE08578;
 XX
 DT 19-NOV-2001 (first entry)
 XX
 DE Human serum albumin (HSA).
 XX
 XX Human: albumin; cancer; cell proliferation; drug screening; biopsy.
 XX
 KW Homo sapiens.
 XX
 OS US6274305-B1.
 XX
 PN 14-AUG-2001.
 PD 19-DEC-1996; 96US-0769746.
 PF 19-DEC-1996; 96US-0769746.
 PR 19-DEC-1996; 96US-0769746.
 XX
 PA (TUFT) UNIV TUFTS.
 XX
 PI Sonenschein C, Soto AM;
 XX
 DR WPI: 2001-540371/60.
 DR N-PSDB; AAD11488.
 XX
 PT Measuring human cell proliferation, useful in drug screening to
 PT determine the potential for inhibiting cancer cell proliferation and
 PT for evaluating biopsied tumors, comprises employing albumin-derived
 PT peptide -
 XX

PS Claim 1; Fig 1; 20pp; English.

XX The invention related to a method for testing cancer cells. The method is
CC useful for measuring human cancer cell proliferation, particularly for
CC determining the potential for inhibiting cancer cells proliferation using
CC albumin-derived peptides. The invention is also useful for drug screening
CC assays, as well as for evaluating biopsied tumours. The present sequence
CC is human serum albumin (HSA) related to the invention.

XX Sequence 585 AA:

Query Match 100.0%; Score 3103; DB 22; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGGENFKALVLAFAOYLQCCPEDHVKLVNTEYERAKTCVADESAB 60
DB 1 DAHSEVAHFRKDLGGENFKALVLAFAOYLQCCPEDHVKLVNTEYERAKTCVADESAB 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHDDNPNLPRLVREPV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHDDNPNLPRLVREPV 120
QY 121 DVWCTAFHNEETFLKLYEIAARRHPYFAPELLFFAKRYKAFTCCOAAADKACLLP 180
DB 121 DVWCTAFHNEETFLKLYEIAARRHPYFAPELLFFAKRYKAFTCCOAAADKACLLP 180
QY 181 KIDELRDEGKASSAKORLKASLOKFGERAFAKMAVARLSORPPKAEFAVSKLYVDLTK 240
DB 181 KIDELRDEGKASSAKORLKASLOKFGERAFAKMAVARLSORPPKAEFAVSKLYVDLTK 240
QY 241 VHTTECHGDLLECCADRADLAKYICENODSISSKLECCCKPLLEKSHCIAEYENDMPA 300
DB 241 VHTTECHGDLLECCADRADLAKYICENODSISSKLECCCKPLLEKSHCIAEYENDMPA 300
QY 301 DLPSTLAADPVESKDVCKNVAEAKDVLGMFLYIARHPDYSVLLRLAKYETTLK 360
DB 301 DLPSTLAADPVESKDVCKNVAEAKDVLGMFLYIARHPDYSVLLRLAKYETTLK 360
QY 361 CAAADHCEYAKYVDFEKPFLVEBPONLIKONCELFOLGKYKONALLVRYTKVQVST 420
DB 361 CAAADHCEYAKYVDFEKPFLVEBPONLIKONCELFOLGKYKONALLVRYTKVQVST 420
QY 421 PTLVEVSRLNGKVGSCCKHPKAMPKCAEDYLSVNLQICVHEKTPVSDRATKCTTES 480
DB 421 PTLVEVSRLNGKVGSCCKHPKAMPKCAEDYLSVNLQICVHEKTPVSDRATKCTTES 480
QY 481 LVNRRCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKQOTALVELYKHKPKAT 540
DB 481 LVNRRCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKQOTALVELYKHKPKAT 540
QY 541 KEOLKAVMDFAAFVCKCKADKKECTCFAEKGGKLVAAASQALGL 585
DB 541 KEOLKAVMDFAAFVCKCKADKKECTCFAEKGGKLVAAASQALGL 585

RESULT 14

ID. ABG63321 standard; protein; 585 AA.

XX ABG63321;

DT 27-AUG-2002 (first entry)

XX Human serum albumin (HSA) protein.

KM albumin fusion protein; therapeutic protein X; human albumin; HA;

KM human serum albumin; HSA; cancer; reproductive disorder;

KM digestive disorder; immune disorder; endocrine disorder;

KM haematopoietic disorder; neural disorder; connective disorder;

KM cytostatic; anti-infectivity; anti-inflammatory; anticancer;

KM immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;

KM neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;

KM osteopathic; antiarthritic.

XX Homo sapiens.

XX WO200177137-A1.

XX 18-OCT-2001.

XX 12-APR-2001; 2001WO-US11988.

XX 12-APR-2000; 2000US-229358P.

XX 25-APR-2000; 2000US-199384P.

XX 21-DEC-2000; 2000US-256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Haseltine WA;

XX WPI; 2002-010886/01.

XX N-PSDB; ABK93280.

XX New fusion protein for treating disease e.g. diabetes comprises an

XX albumin fused to a therapeutic protein -

XX Claim 1; Fig 15; 2102pp; English.

CC The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or
CC disorder that may be modulated by therapeutic protein X. The albumin
CC extends the shelf-life of protein X, and may increase its biological
CC in vitro/in vivo activity. The protein is useful for treating and
CC diagnosing disorders such as cancer, reproductive disorders, digestive
CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
CC (e.g. diabetes), haematopoietic disorders, neural disorders
CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
CC (e.g. osteoporosis, arthritis). The present sequence represents HSA
CC (HA) protein.

XX Sequence 585 AA:

Query Match 100.0%; Score 3103; DB 23; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGGENFKALVLAFAOYLQCCPEDHVKLVNTEYERAKTCVADESAB 60
DB 1 DAHSEVAHFRKDLGGENFKALVLAFAOYLQCCPEDHVKLVNTEYERAKTCVADESAB 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHDDNPNLPRLVREPV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHDDNPNLPRLVREPV 120
QY 121 DVWCTAFHNEETFLKLYEIAARRHPYFAPELLFFAKRYKAFTCCOAAADKACLLP 180
DB 121 DVWCTAFHNEETFLKLYEIAARRHPYFAPELLFFAKRYKAFTCCOAAADKACLLP 180
QY 181 KIDELRDEGKASSAKORLKASLOKFGERAFAKMAVARLSORPPKAEFAVSKLYVDLTK 240
DB 181 KIDELRDEGKASSAKORLKASLOKFGERAFAKMAVARLSORPPKAEFAVSKLYVDLTK 240
QY 241 VHTTECHGDLLECCADRADLAKYICENODSISSKLECCCKPLLEKSHCIAEYENDMPA 300
DB 241 VHTTECHGDLLECCADRADLAKYICENODSISSKLECCCKPLLEKSHCIAEYENDMPA 300
QY 301 DLPSTLAADPVESKDVCKNVAEAKDVLGMFLYIARHPDYSVLLRLAKYETTLK 360
DB 301 DLPSTLAADPVESKDVCKNVAEAKDVLGMFLYIARHPDYSVLLRLAKYETTLK 360
QY 361 CAAADHCEYAKYVDFEKPFLVEBPONLIKONCELFOLGKYKONALLVRYTKVQVST 420
DB 361 CAAADHCEYAKYVDFEKPFLVEBPONLIKONCELFOLGKYKONALLVRYTKVQVST 420

Db 361 CAADPHCYAKVFDEKPLVEEPONLIKONCELFEOLGEEKFONALIVRTKKVPQVST 420
 QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCTES 480
 Db 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCTES 480
 QY 481 LVNRRPCSALEVEDETVPKFEFNAETFFHADICTLSKEKROIKQTALVELVHKRPAT 540
 Db 481 LVNRRPCSALEVEDETVPKFEFNAETFFHADICTLSKEKROIKQTALVELVHKRPAT 540
 QY 541 KEOLKAVMDPFAAFVEKCKKADKETCFAEKGKTLVAASQAALGI 585
 Db 541 KEOLKAVMDPFAAFVEKCKKADKETCFAEKGKTLVAASQAALGI 585

RESULT 15

ABJ00986
 ID ABJ00986 standard; Protein: 585 AA.

AC ABJ00986;

DT 05-SEP-2002 (first entry)

XX B lymphocyte stimulator protein binding peptide related protein.

DE B lymphocyte stimulator protein binding protein: Blys; immune disease;
 XX allergy; proliferative disease; infectious disease; arteriosclerosis;
 KW inflammatory disorder; hypergammaglobulinaemia; blood clotting;
 KW ischaemia; graft-versus-host disease; neurodegenerative disease;
 KW immunosuppressive; nephrotropic; antirheumatic; antithyroid;
 KW neutropenic; cytostatic; immunostimulant; antitumor; anti-HIV;
 KW antiaesthetic; antiallergic; thyromimetic; antianemic; haemostatic;
 KW dermatological; antiinflammatory; cardiac; ophthalmological; uropathic;
 KW antidiabetic; antithyroid; antidepressant; hepatotropic.

XX Homo sapiens.

XX MO200216411-A2.

XX 28-FEB-2002.

XX 17-AUG-2001; 2001WO-US25850.

XX 18-AUG-2000; 2000US-226700P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Beltzer JP, Potter DM, Fleming TL, Rosen CA;

XX WPI; 2002-499775/53.

XX The treatment of various diseases e.g. rheumatoid arthritis, comprises

XX administering B lymphocyte stimulator binding polypeptide

XX Disclosure: Page 379-382; 387pp; English.

XX The present invention relates to the treatment, prevention or
 CC amelioration of a disease or disorder associated with: aberrant B
 CC lymphocyte stimulator (Blys), Blys receptor expression or activity; cells
 CC of haematopoietic origin; or proliferative disease; and reducing,
 CC inhibiting or stimulating immunoglobulin production, B cell proliferation
 CC and graft rejection involving administration of Blys binding polypeptide.
 CC The Blys binding polypeptides are used in the treatment, prevention or
 CC amelioration of diseases such as immune system diseases, proliferative
 CC diseases, diseases of cells of hematopoietic origin, graft rejection,
 CC allergies, infectious diseases, arteriosclerosis, inflammatory disorders,
 CC hypergammaglobulinaemia, blood clotting disorders, ischaemia, and
 CC neurodegenerative diseases. The present sequence is a protein described
 CC in the invention.

XX Sequence 585 AA:

Query Match 100.0%; Score 3103; DB 23; Length 585;

Best Local Similarity 100.0%; Pred. No. 1e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRKRDGSENFKALVIAFAOYLQOCPFEDHVKLVNEVEFAKTCVADSAAE 60
 Db 1 DAHSEVAHRKRDGSENFKALVIAFAOYLQOCPFEDHVKLVNEVEFAKTCVADSAAE 60
 QY 61 NCDKSLHTLFEGDKLCTVAATLRETYGEMADCCAKOPEPNECFLOHKDNPRLVPREV 120
 Db 61 NCDKSLHTLFEGDKLCTVAATLRETYGEMADCCAKOPEPNECFLOHKDNPRLVPREV 120
 QY 121 DVMTAFHDNEETFLKYLVEIARRHPYFAPELLFFAKRYKAFTCCOADAACCLP 180
 Db 121 DVMTAFHDNEETFLKYLVEIARRHPYFAPELLFFAKRYKAFTCCOADAACCLP 180
 QY 181 KLDELRDGKASAKORLKASLOKFGERRAKANAVARLSORFPAEFAEYSKLVTDLTJK 240
 Db 181 KLDELRDGKASAKORLKASLOKFGERRAKANAVARLSORFPAEFAEYSKLVTDLTJK 240
 QY 241 VHTCCGHDDLCEADDRADLAKYICENODSISKLKCECEKPLLEKSHCIAEVENDEMPA 300
 Db 241 VHTCCGHDDLCEADDRADLAKYICENODSISKLKCECEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLADPFAVSKVCKNYAKADVFLGMFLYETARRHPDYSVLLRLAKTYETLLEKC 360
 Db 301 DLPSLADPFAVSKVCKNYAKADVFLGMFLYETARRHPDYSVLLRLAKTYETLLEKC 360
 QY 361 CAADPHCYAKVFDEKPLVEEPONLIKONCELFEOLGEEKFONALIVRTKKVPQVST 420
 Db 361 CAADPHCYAKVFDEKPLVEEPONLIKONCELFEOLGEEKFONALIVRTKKVPQVST 420
 QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCTES 480
 Db 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRVTKCTES 480
 QY 481 LVNRRPCSALEVEDETVPKFEFNAETFFHADICTLSKEKROIKQTALVELVHKRPAT 540
 Db 481 LVNRRPCSALEVEDETVPKFEFNAETFFHADICTLSKEKROIKQTALVELVHKRPAT 540
 QY 541 KEOLKAVMDPFAAFVEKCKKADKETCFAEKGKTLVAASQAALGI 585
 Db 541 KEOLKAVMDPFAAFVEKCKKADKETCFAEKGKTLVAASQAALGI 585

Search completed: July 22, 2003, 11:43:43

Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 22, 2003, 11:43:03 ; Search time 18 Seconds
(without alignments)
956.245 Million cell updates/sec

Title: US-09-833-118-18
3103
Perfect score: 1 DAHSEVAHREKDLGEENFK.....TCFAEGKKLVAAQALGL 585
Sequence:

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	585	1	US-08-153-799-14 Sequence 14, Appl
2	3103	100.0	585	2	US-08-702-572-2 Sequence 2, Appl
3	3103	100.0	585	4	US-08-769-746-2 Sequence 2, Appl
4	3103	100.0	610	2	US-08-797-689-2 Sequence 2, Appl
5	3103	100.0	783	1	US-08-256-938-2 Sequence 2, Appl
6	3103	100.0	787	1	US-08-256-938-4 Sequence 4, Appl
7	3103	100.0	787	2	US-08-797-689-16 Sequence 16, Appl
8	3099	99.9	609	1	US-08-222-619-3 Sequence 3, Appl
9	3099	99.9	609	1	US-08-433-037-4 Sequence 4, Appl
10	3099	99.9	609	5	US-08-897-956A-2 Sequence 2, Appl
11	3095	99.9	978	4	US-08-897-956A-3 Sequence 3, Appl
12	3095	99.7	978	4	US-08-897-956A-3 Sequence 3, Appl
13	3093	99.7	585	1	US-08-448-196A-3 Sequence 3, Appl
14	3093	99.7	585	2	US-08-448-196A-3 Sequence 3, Appl
15	2458.5	79.2	583	1	US-08-984-176-1 Sequence 1, Appl
16	2450.5	79.0	583	1	US-08-448-196A-5 Sequence 5, Appl
17	2432.5	78.4	583	1	US-08-448-196A-4 Sequence 4, Appl
18	2426	78.2	584	1	US-08-448-196A-6 Sequence 6, Appl
19	2389	77.0	582	1	US-08-448-196A-7 Sequence 7, Appl
20	1249.5	40.3	609	1	US-08-134-638-1 Sequence 1, Appl
21	1249.5	40.3	609	5	US-08-222-619-4 Sequence 4, Appl
22	1206.5	38.9	590	5	PCT-US95-04075-4 Sequence 4, Appl
23	1206.5	38.9	590	4	US-08-377-309-2 Sequence 2, Appl
24	1206.5	38.9	590	4	US-08-505-012-5 Sequence 2, Appl
25	1206.5	38.9	590	4	US-08-505-012-5 Sequence 5, Appl
26	1206.5	38.9	590	5	US-09-186-949A-3 Sequence 3, Appl
27	1206.5	38.9	609	4	PCT-US96-00996-5 Sequence 5, Appl
					US-09-186-949A-2 Sequence 2, Appl

28	1164.5	37.5	579	1	US-08-448-196A-8	Sequence 8, Appl
29	1055	34.0	599	1	US-08-222-619-2	Sequence 2, Appl
30	1055	34.0	599	4	US-08-221-767-24	Sequence 24, Appl
31	1055	34.0	599	5	PCT-US95-04075-2	Sequence 2, Appl
32	926	29.8	393	2	US-08-377-309-7	Sequence 7, Appl
33	926	29.8	393	4	US-09-186-723-7	Sequence 7, Appl
34	926	29.8	393	4	US-08-505-012-10	Sequence 10, Appl
35	926	29.8	393	4	US-08-505-012-10	Sequence 10, Appl
36	926	29.8	393	5	US-09-186-949A-8	Sequence 8, Appl
37	926	29.8	393	5	PCT-US96-00996-10	Sequence 10, Appl
38	777	25.0	324	4	US-08-505-012-12	Sequence 12, Appl
39	777	25.0	325	2	PCT-US96-00996-12	Sequence 12, Appl
40	777	25.0	325	4	US-08-377-309-8	Sequence 8, Appl
41	777	25.0	325	4	US-09-186-723-8	Sequence 8, Appl
42	777	25.0	325	4	US-08-505-012-11	Sequence 11, Appl
43	777	25.0	325	4	US-09-186-949A-9	Sequence 9, Appl
44	747.5	24.1	325	5	PCT-US96-00996-11	Sequence 11, Appl
45	684.5	22.1	389	2	US-08-448-196A-9	Sequence 9, Appl
					US-08-377-309-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-153-799-14
Sequence 14, Application US/08153799
Patent No. 5766883
GENERAL INFORMATION:
APPLICANT: Ballance, David J
APPLICANT: Goodey, Andrew R
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: R Hain Swope, BOC Health Care Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill
STATE: New Jersey
COUNTRY: USA
ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear

```
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Region
LOCATION: 369..419 /note="Alternative C-termini of
OTHER INFORMATION: HSA(1-n)"
FEATURE:
NAME/KEY: Region
LOCATION: 1..585
OTHER INFORMATION: /note="Amino acid sequence of
OTHER INFORMATION: natural HSA"
US-08-153-799-14

Query Match 100.0%; Score 3103; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.6e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRFKDGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSAAE 60
DB 1 DAHSEVAHRRFKDGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSAAE 60
QY 61 NCDSLHTLFEGDKICTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPLVPREV 120
DB 61 NCDSLHTLFEGDKICTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPLVPREV 120
QY 121 DVNCTAFHDNEETFLKKYLYEIARRHRYFYAPPELLFFAKRYKAFTCCQAADKAACLLP 180
DB 121 DVNCTAFHDNEETFLKKYLYEIARRHRYFYAPPELLFFAKRYKAFTCCQAADKAACLLP 180
QY 181 KLDELREGKASSAKORLKCSLOKFGBRAFKAAVAVRLSORPFAKAEVSKLVTDLTK 240
DB 181 KLDELREGKASSAKORLKCSLOKFGBRAFKAAVAVRLSORPFAKAEVSKLVTDLTK 240
QY 241 VHTCCGHDLLCEADDDRADLAKYICENODSISSKLKECCPEPLLEKSHCIAEVNDMPA 300
DB 241 VHTCCGHDLLCEADDDRADLAKYICENODSISSKLKECCPEPLLEKSHCIAEVNDMPA 300
QY 301 DLPSLADFEVSKVCKNYAEAKDVFGLMFLYEFARRHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLADFEVSKVCKNYAEAKDVFGLMFLYEFARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHECYAKVDEFRPLVEEPONLIKONCELFEOLGEEKFONALLVRYTKKVPYST 420
DB 361 CAADPHECYAKVDEFRPLVEEPONLIKONCELFEOLGEEKFONALLVRYTKKVPYST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPFSALVEVETVPKFEFNAETFTFHADICTLSKEKROIKKOTALVELVHKPKAT 540
DB 481 LVNRRPFSALVEVETVPKFEFNAETFTFHADICTLSKEKROIKKOTALVELVHKPKAT 540
QY 541 KEQLKAVWDPAFAFVEKCKKADDEKTCFAEEGKRLVAASQAALGI 585
DB 541 KEQLKAVWDPAFAFVEKCKKADDEKTCFAEEGKRLVAASQAALGI 585

RESULT 2
US-08-702-572-2
Sequence 2, Application US/08702572
Patent No. 5965386
GENERAL INFORMATION:
APPLICANT: Kerry-Williams, Sean M
APPLICANT: Gilbert, Sarah C
TITLE OF INVENTION: Yeast Strains and Modified Albumins
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Centeon L.L.C.
STREET: 1020 First Avenue
```

```
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-1310
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,572
FILING DATE: 11-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 95/23857
FILING DATE: 1-MAR-1995
APPLICATION NUMBER: GB 9404270.2
FILING DATE: 5-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Naomi Biswas
REGISTRATION NUMBER: 38,384
REFERENCE/DOCKET NUMBER: CE0114 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610/878/4294
TELEFAX: 610/878/4221
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-702-572-2

Query Match 100.0%; Score 3103; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.6e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRFKDGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSAAE 60
DB 1 DAHSEVAHRRFKDGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADSAAE 60
QY 61 NCDSLHTLFEGDKICTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPLVPREV 120
DB 61 NCDSLHTLFEGDKICTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPLVPREV 120
QY 121 DVNCTAFHDNEETFLKKYLYEIARRHRYFYAPPELLFFAKRYKAFTCCQAADKAACLLP 180
DB 121 DVNCTAFHDNEETFLKKYLYEIARRHRYFYAPPELLFFAKRYKAFTCCQAADKAACLLP 180
QY 181 KLDELREGKASSAKORLKCSLOKFGBRAFKAAVAVRLSORPFAKAEVSKLVTDLTK 240
DB 181 KLDELREGKASSAKORLKCSLOKFGBRAFKAAVAVRLSORPFAKAEVSKLVTDLTK 240
QY 241 VHTCCGHDLLCEADDDRADLAKYICENODSISSKLKECCPEPLLEKSHCIAEVNDMPA 300
DB 241 VHTCCGHDLLCEADDDRADLAKYICENODSISSKLKECCPEPLLEKSHCIAEVNDMPA 300
QY 301 DLPSLADFEVSKVCKNYAEAKDVFGLMFLYEFARRHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLADFEVSKVCKNYAEAKDVFGLMFLYEFARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHECYAKVDEFRPLVEEPONLIKONCELFEOLGEEKFONALLVRYTKKVPYST 420
DB 361 CAADPHECYAKVDEFRPLVEEPONLIKONCELFEOLGEEKFONALLVRYTKKVPYST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPFSALVEVETVPKFEFNAETFTFHADICTLSKEKROIKKOTALVELVHKPKAT 540
DB 481 LVNRRPFSALVEVETVPKFEFNAETFTFHADICTLSKEKROIKKOTALVELVHKPKAT 540
```

QY 541 KEOLKAYMDPFAAFVEKCKKADKDKETCFABEGKKLVASQAALGL 585
Db 541 KEOLKAYMDPFAAFVEKCKKADKDKETCFABEGKKLVASQAALGL 585

RESULT 3

US-08-769-746-2
Sequence 2, Application US/08769746
Patent No. 6274305
GENERAL INFORMATION:
APPLICANT: Sonnenschein, Carlos
APPLICANT: Soto, Ana M.
TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,746
FILING DATE: 19-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: MBRT-02584
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-769-746-2

Query Match 100.0%; Score 3103; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.6e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHHRFKDGEENFKALVLIARAOYLQCCPFEDHVKLVNEVEFAKTCVADDSAE 60
Db 1 DAHSEVAHHRFKDGEENFKALVLIARAOYLQCCPFEDHVKLVNEVEFAKTCVADDSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKOEPERNECFLOHKDNPMLPRLVPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGENADCCAKOEPERNECFLOHKDNPMLPRLVPEV 120
QY 121 DVNCTAHNDNEFTLKLYLEIARRHRYFAPELLEFPAKRYKAAFTCCQAADKAACILP 180
Db 121 DVNCTAHNDNEFTLKLYLEIARRHRYFAPELLEFPAKRYKAAFTCCQAADKAACILP 180
QY 181 KLDELREGKASSAKORLKASLOKGEERAFKAAVAVARLSORPKAEFAEYSKLVTLTK 240
Db 181 KLDELREGKASSAKORLKASLOKGEERAFKAAVAVARLSORPKAEFAEYSKLVTLTK 240
QY 241 VHEECGHDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHEECGHDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
QY 301 DPLSLADPYESKVCNRYAKAVFLGMFLYEVARRHPDYSVLLRLAKTYTTLEK 360
Db 301 DPLSLADPYESKVCNRYAKAVFLGMFLYEVARRHPDYSVLLRLAKTYTTLEK 360

QY 361 CAADPHCEYAKVDEKPLVEEPONLIKONCELEPOLGEXKFORALLVRYTKKPOVST 420
Db 361 CAADPHCEYAKVDEKPLVEEPONLIKONCELEPOLGEXKFORALLVRYTKKPOVST 420
QY 421 PTLVEVSRLGKVKSCCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
Db 421 PTLVEVSRLGKVKSCCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPFSALVEDETVPKFEFNAETFTPHADITLTSEKEKQIKOTALVELVNHKRYAT 540
Db 481 LVNRRPFSALVEDETVPKFEFNAETFTPHADITLTSEKEKQIKOTALVELVNHKRYAT 540
QY 541 KEOLKAYMDPFAAFVEKCKKADKDKETCFABEGKKLVASQAALGL 585
Db 541 KEOLKAYMDPFAAFVEKCKKADKDKETCFABEGKKLVASQAALGL 585

RESULT 4

US-08-797-689-2
Sequence 2, Application US/08797689
Patent No. 5876969
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
APPLICANT: Fournier, Alain
APPLICANT: Guilton, Jean-Dominique
APPLICANT: Jung, Gerard
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION DATA:
PCT/FR93/00085
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-2

Query Match 100.0%; Score 3103; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 7e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 DAHSEVAHRRKDLGEENFKALVLIAPAOYLQOCPEFDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHSEVAHRRKDLGEENFKALVLIAPAOYLQOCPEFDHVKLVNEVTEFAKTCVADESAE 84
OY 61 NCDSLHRLFGDKLCTVATLRETYGEMADCCAKOEPENECFLOHKDNPMLPRLVREY 120
DB 85 NCDSLHRLFGDKLCTVATLRETYGEMADCCAKOEPENECFLOHKDNPMLPRLVREY 144
OY 121 DVMCTAFHDNEETFLKLYEIAARRHPYFVAPBELLFFAKRYKAFTBCCQAADKACLLP 180
DB 145 DVMCTAFHDNEETFLKLYEIAARRHPYFVAPBELLFFAKRYKAFTBCCQAADKACLLP 204
OY 181 KLDELROEGKASSAKORLKASLOKFGERRAKAMAVARLSORFPAEFAEYSKLVTDLTK 240
DB 205 KLDELROEGKASSAKORLKASLOKFGERRAKAMAVARLSORFPAEFAEYSKLVTDLTK 264
OY 241 VHTECCHDILLECADRDADLAKYICENDSISSKIECCERPLEKSHCIAEVENDEMPA 300
DB 265 VHTECCHDILLECADRDADLAKYICENDSISSKIECCERPLEKSHCIAEVENDEMPA 324
OY 301 DLPSIADPFVESKDVCKNYAEKDVFLGMPLEYARRRDPDYSVLLRLAKTYETTLK 360
DB 325 DLPSIADPFVESKDVCKNYAEKDVFLGMPLEYARRRDPDYSVLLRLAKTYETTLK 384
OY 361 CAADPHCEYAKVDEFRPLVEEPONLIKONCELEQGEYKFOALLVRYTKKVPYST 420
DB 385 CAADPHCEYAKVDEFRPLVEEPONLIKONCELEQGEYKFOALLVRYTKKVPYST 444
OY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRYTKCTES 480
DB 445 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRYTKCTES 504
OY 481 LVNRRPCSALEVDITYPKPENAETFFHADICTLSKERQIKKOTLALVELVKKPRAT 540
DB 505 LVNRRPCSALEVDITYPKPENAETFFHADICTLSKERQIKKOTLALVELVKKPRAT 564
OY 541 KEQLKAVMDPFAAFYEKCKKADDKETCFABEGKRLVAASQAALGL 585
DB 565 KEQLKAVMDPFAAFYEKCKKADDKETCFABEGKRLVAASQAALGL 609

```

RESULT 5
 US-08-256-938-2
 ; Sequence 2, Application US/08256938
 ; Patent No. 5665863
 ; GENERAL INFORMATION:
 ; APPLICANT: Yeh, Patrice
 ; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
 ; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.
 ; STREET: 500 Arcola Road, 3c43
 ; CITY: Collegeville
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19426
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: System 7.1
 ; SOFTWARE: Word 5.0 (Patentin)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/256,938
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: FR 92/01065
 ; FILING DATE: 31-JAN-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Goodman, Rosanne

```

; REGISTRATION NUMBER: 32,534
; REFERENCE/DOCKET NUMBER: ST92007-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3817
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 783 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-938-2

Query Match      100.0%; Score 3103; DB 1; Length 783;
Best Local Similarity 100.0%; Pred. No. 1e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 DAHSEVAHRRKDLGEENFKALVLIAPAOYLQOCPEFDHVKLVNEVTEFAKTCVADESAE 60
DB 25 DAHSEVAHRRKDLGEENFKALVLIAPAOYLQOCPEFDHVKLVNEVTEFAKTCVADESAE 84
OY 61 NCDSLHRLFGDKLCTVATLRETYGEMADCCAKOEPENECFLOHKDNPMLPRLVREY 120
DB 85 NCDSLHRLFGDKLCTVATLRETYGEMADCCAKOEPENECFLOHKDNPMLPRLVREY 144
OY 121 DVMCTAFHDNEETFLKLYEIAARRHPYFVAPBELLFFAKRYKAFTBCCQAADKACLLP 180
DB 145 DVMCTAFHDNEETFLKLYEIAARRHPYFVAPBELLFFAKRYKAFTBCCQAADKACLLP 204
OY 181 KLDELROEGKASSAKORLKASLOKFGERRAKAMAVARLSORFPAEFAEYSKLVTDLTK 240
DB 205 KLDELROEGKASSAKORLKASLOKFGERRAKAMAVARLSORFPAEFAEYSKLVTDLTK 264
OY 241 VHTECCHDILLECADRDADLAKYICENDSISSKIECCERPLEKSHCIAEVENDEMPA 300
DB 265 VHTECCHDILLECADRDADLAKYICENDSISSKIECCERPLEKSHCIAEVENDEMPA 324
OY 301 DLPSIADPFVESKDVCKNYAEKDVFLGMPLEYARRRDPDYSVLLRLAKTYETTLK 360
DB 325 DLPSIADPFVESKDVCKNYAEKDVFLGMPLEYARRRDPDYSVLLRLAKTYETTLK 384
OY 361 CAADPHCEYAKVDEFRPLVEEPONLIKONCELEQGEYKFOALLVRYTKKVPYST 420
DB 385 CAADPHCEYAKVDEFRPLVEEPONLIKONCELEQGEYKFOALLVRYTKKVPYST 444
OY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRYTKCTES 480
DB 445 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTPVSDRYTKCTES 504
OY 481 LVNRRPCSALEVDITYPKPENAETFFHADICTLSKERQIKKOTLALVELVKKPRAT 540
DB 505 LVNRRPCSALEVDITYPKPENAETFFHADICTLSKERQIKKOTLALVELVKKPRAT 564
OY 541 KEQLKAVMDPFAAFYEKCKKADDKETCFABEGKRLVAASQAALGL 585
DB 565 KEQLKAVMDPFAAFYEKCKKADDKETCFABEGKRLVAASQAALGL 609

```

RESULT 6
 US-08-256-938-4
 ; Sequence 4, Application US/08256938
 ; Patent No. 5665863
 ; GENERAL INFORMATION:
 ; APPLICANT: Yeh, Patrice
 ; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
 ; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.
 ; STREET: 500 Arcola Road, 3c43
 ; CITY: Collegeville
 ; STATE: PA

COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.0 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,938
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/01065
FILING DATE: 31-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Rosanne
REGISTRATION NUMBER: 32,534
REFERENCE/DOCKET NUMBER: ST92007-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3817
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 787 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-256-938-4

Query Match 100.0%; Score 3103; DB 1; Length 787;
Best Local Similarity 100.0%; Pred. No. 1e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGGEENFKALVLIAPQYLQCCPEFDHVKLVNEVTEFAKTCVADSAE 60
DB 203 DAHSEVAHFRKDLGGEENFKALVLIAPQYLQCCPEFDHVKLVNEVTEFAKTCVADSAE 262
QY 61 NCCKSLHTLFGDKICTVATLRETYGEMADCCAKOEPERNCFLOHKDNDNLPRLVREY 120
DB 263 NCCKSLHTLFGDKICTVATLRETYGEMADCCAKOEPERNCFLOHKDNDNLPRLVREY 322
QY 121 DVNCTAFHNDNEFTLKKYLYEIAIRRHRYFYAPPELLFFAKRYKAAFTCCOAAKACLLP 180
DB 323 DVNCTAFHNDNEFTLKKYLYEIAIRRHRYFYAPPELLFFAKRYKAAFTCCOAAKACLLP 362
QY 181 KLDELDEGKASSAKORLKACASLOKFGERAFAKMAVARLSQRPKAEFAVSKLVYDITK 240
DB 383 KLDELDEGKASSAKORLKACASLOKFGERAFAKMAVARLSQRPKAEFAVSKLVYDITK 442
QY 241 VHTCCGHDLLECADDRAADAKYICENODSISIKLBCCCKPILKSHCAEYENDMPA 300
DB 443 VHTCCGHDLLECADDRAADAKYICENODSISIKLBCCCKPILKSHCAEYENDMPA 502
QY 301 DLPSLADEYVESKDVCKNYAEAKDVFLGMFLYEYARRHPYVYLLRLAKTYETTLK 360
DB 503 DLPSLADEYVESKDVCKNYAEAKDVFLGMFLYEYARRHPYVYLLRLAKTYETTLK 562
QY 361 CAADPHCEYAKVYDEFKPLVEEPONLIKONCELFEOLGSEYKFONALLVYTKKVPVST 420
DB 563 CAADPHCEYAKVYDEFKPLVEEPONLIKONCELFEOLGSEYKFONALLVYTKKVPVST 622
QY 421 PTLVEYSRNIGKYGSCCKHPEAKRMPACADYLSVYLNOLCVLHEKTPVSDRYTKCTES 480
DB 623 PTLVEYSRNIGKYGSCCKHPEAKRMPACADYLSVYLNOLCVLHEKTPVSDRYTKCTES 682
QY 481 LVNRRPCFSALVEDEVYKPEFNAETFTFHADICTLSEKEROIKKOTALVELYKHKPKAT 540
DB 683 LVNRRPCFSALVEDEVYKPEFNAETFTFHADICTLSEKEROIKKOTALVELYKHKPKAT 742
QY 541 KEOLKAVMDFAAFVEKCKADK ETCFAEGKKLVAAQAALGL 585
DB 743 KEOLKAVMDFAAFVEKCKADK ETCFAEGKKLVAAQAALGL 787

RESULT 7

US-08-797-689-16
Sequence 16, Application US/08797689
Patent No. 5876969
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
APPLICANT: Fournier, Alain
APPLICANT: Gultton, Jean-Dominique
APPLICANT: Jung, Gerard
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 787 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-16

Query Match 100.0%; Score 3103; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. 1e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGGEENFKALVLIAPQYLQCCPEFDHVKLVNEVTEFAKTCVADSAE 60
DB 203 DAHSEVAHFRKDLGGEENFKALVLIAPQYLQCCPEFDHVKLVNEVTEFAKTCVADSAE 262
QY 61 NCCKSLHTLFGDKICTVATLRETYGEMADCCAKOEPERNCFLOHKDNDNLPRLVREY 120
DB 263 NCCKSLHTLFGDKICTVATLRETYGEMADCCAKOEPERNCFLOHKDNDNLPRLVREY 322
QY 121 DVNCTAFHNDNEFTLKKYLYEIAIRRHRYFYAPPELLFFAKRYKAAFTCCOAAKACLLP 180
DB 323 DVNCTAFHNDNEFTLKKYLYEIAIRRHRYFYAPPELLFFAKRYKAAFTCCOAAKACLLP 362
QY 181 KLDELDEGKASSAKORLKACASLOKFGERAFAKMAVARLSQRPKAEFAVSKLVYDITK 240
DB 383 KLDELDEGKASSAKORLKACASLOKFGERAFAKMAVARLSQRPKAEFAVSKLVYDITK 442

Oy	241	YHTECHDHLIECCADRDADLAKYICENODSISSKLKECKECPLEKSHCIAEVENDEMPA	300
Dd	443	YHTECHGHDLECCADRDADLAKYICENODSISSKLKECKECPLEKSHCIAEVENDEMPA	502
Oy	301	DLPSLADPVEBKDYCKNYAEAKDYFLGMFLYEYARRHPDYSVLLLRILAKTYETLLEKC	360
Dd	503	DLPSLADPVEBKDYCKNYAEAKDYFLGMFLYEYARRHPDYSVLLLRILAKTYETLLEKC	562
Oy	361	CAAPDHBCYAUVFDEBFKPLVEEPONLIKONCELPOLGEYKONALLVRYTKKPOVST	420
Dd	563	CAAPDHBCYAUVFDEBFKPLVEEPONLIKONCELPOLGEYKONALLVRYTKKPOVST	622
Oy	421	PLVYVSURLKVGSKCKCHHPAKRMPCAEDYLSVLLNOLCVLHEKTPVSDRYTKCCTES	480
Dd	623	PLVYVSURLKVGSKCKCHHPAKRMPCAEDYLSVLLNOLCVLHEKTPVSDRYTKCCTES	682
Oy	481	LNNRPPCSALEVDETYVPEKERNAEFTTHADICLTSEKERQIKKOTALVELVHKRPAT	540
Dd	683	LNNRPPCSALEVDETYVPEKERNAEFTTHADICLTSEKERQIKKOTALVELVHKRPAT	742
Oy	541	KEOLKAUVMDDFAAFYEKCKKADDKETCFMEEGKSLVAAQSAALGL	585
Dd	743	KEOLKAUVMDDFAAFYEKCKKADDKETCFMEEGKSLVAAQSAALGL	787

RESULT 8
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352

1 GENERAL INFORMATION:
2 APPLICANT: Lichenstein, Henri
3 APPLICANT: Lyons, David
4 APPLICANT: Wurfel, Mark
5 APPLICANT: Wighn, Samuel
6 TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
7 TITLE OF INVENTION: Protein
8 NUMBER OF SEQUENCES: 33
9 CORRESPONDENCE ADDRESS:
10 ADDRESSEE: Amgen Center, Patent Operations/RRC
11 STREET: 1840 Dehavilland Drive
12 CITY: Thousand Oaks
13 STATE: California
14 COUNTRY: U.S.
15 ZIP: 91320-1789
16
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: FLOPPY disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: Patent In Release #1.0, Version #1.25
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/08/222,619
24 FILING DATE:
25 CLASSIFICATION: 435
26 INFORMATION FOR SEQ ID NO: 3:
27 SEQUENCE CHARACTERISTICS:
28 LENGTH: 609 amino acids
29 TYPE: amino acid
30 STRANDEDNESS: unknown
31 TOPOLOGY: unknown
32 MOLECULE TYPE: protein
33 US-08-222-619-3

Query Match	99.9%	Score 3099	DB 1	Length 609
Best Local Similarity	99.8%	Pred. No. 1.7e-286		
Matches 384	Conservative	1	Mismatches 0	Indels 0
				Gaps 0

QY 1 DAHKSEVAHREFKDLGSEENKALVLAFAQYLQCCPEDHYKLVNTEATFAKTCVADSAE 60
 |||||
 Db 25 DAHKSEVAHREFKDLGSEENKALVLAFAQYLQCCPEDHYKLVNTEATFAKTCVADSAE 84
 |||||
 QY 61 NCDKSLHILFGKCLTVAPLRETYGMADCCAKOEERNECFLOHKDDNPNLPILYRPEV 120
 |||||

Db	85	NCOKSIHTLEGGOKLTQVATLURETGYBMADCSAKQORENEBCEFLQHKNDONPRLPLVLPREV	144
OY	121	DWNCSTAFHNEETFLFKLYUETARKNHPRFYAPELLFFAKRTKAFTECCOAAKRAACLLP	180
Db	145	DVNCSTAFHNEETFLFKLYUETARKNHPRFYAPELLFFAKRTKAFTECCOAAKRAACLLP	204
OY	181	KLDELDEDEGKASSAKQRLCASLOKFGERAFFAMAVALSOPRPAKEAEVSKLVTDTLK	240
Db	205	KLDELDEDEGKASSAKQRLCASLOKFGERAFFAMAVALSOPRPAKEAEVSKLVTDTLK	264
OY	241	VHRECHGDILLECADRADLAKYTCENODSISSTKKECCERPLLEKSHICIAEVNDEMPA	300
Db	265	VHRECHGDILLECADRADLAKYTCENODSISSTKKECCERPLLEKSHICIAEVNDEMPA	324
OY	301	DLESLADFEVSEKDVCKNYAENKDYFLGMFYETARRHPDVSUULLRLAKYETTELK	360
Db	325	DLESLADFEVSEKDVCKNYAENKDYFLGMFYETARRHPDVSUULLRLAKYETTELK	384
OY	361	CAAADPHECYAKVDEFKPLVEBPONLIKONCELEFEOJGEBKFPONALLVRYTKKPVOST	420
Db	385	CAAADPHECYAKVDEFKPLVEBPONLIKONCELEFEOJGEBKFPONALLVRYTKKPVOST	444
OY	421	PLTVEVSRNLGKVGSKCHNRPAAKMPCAEDLSVLNQLCVLHNEKTPVSDRYTKCTES	480
Db	445	PLTVEVSRNLGKVGSKCHNRPAAKMPCAEDLSVLNQLCVLHNEKTPVSDRYTKCTES	504
OY	481	LNNRPFCSALEVDETYVPKEFNAAETFTYBHADICTLSKEKQIKQJALVELYKHKFAT	540
Db	505	LNNRPFCSALEVDETYVPKEFNAAETFTYBHADICTLSKEKQIKQJALVELYKHKFAT	564
OY	541	KBLKAAVMDDEFAFVEKCCAKDKDETCFAEBSBKLYVAASQALGI	585
Db	565	KBLKAAVMDDEFAFVEKCCAKDKDETCFAEBSBKLYVAASQALGI	609

RESULT 9
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828

APPLICANT: Sreekrishna, Kottikanyadan
APPLICANT: Bart, Kathryn A.
APPLICANT: Bierley, Russell A.
APPLICANT: Thill, Gregory P.
APPLICANT: Tschopp, Juerg F.
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
TITLE OF INVENTION: PCHIA PASTORIS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scilly, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/433,037
FILING DATE: 03-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digililo, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 91082
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-433-037-4

Query Match 99.9%; Score 3099; DB 1; Length 609;
Best Local Similarity 99.8%; Pred. No. 1,7e-286;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDGEEENFKALVLIAPAOYLQOCPEFDHVKLVNEVEFAKTCVADSAE 60
DB 25 DAHSEVAHFRKDGEEENFKALVLIAPAOYLQOCPEFDHVKLVNEVEFAKTCVADSAE 84
QY 61 NCKSLHTLFGDKICTYATLRETYGEMADCCAKOEPERNECFLQHKDNDNPLPLVPREV 120
DB 85 NCKSLHTLFGDKICTYATLRETYGEMADCCAKOEPERNECFLQHKDNDNPLPLVPREV 144
QY 121 DVNCTAHNDNEFTLKKYLYEIAARRHYFYAPELLFFAKRYKKAFTCCOAAKACLLP 180
DB 145 DVNCTAHNDNEFTLKKYLYEIAARRHYFYAPELLFFAKRYKKAFTCCOAAKACLLP 204
QY 181 KLDELREGKASSAKORLKCASLOKFGERAFAKANAVALRSORPKAEFAEYSKLVTDLT 240
DB 205 KLDELREGKASSAKORLKCASLOKFGERAFAKANAVALRSORPKAEFAEYSKLVTDLT 264
QY 241 VHTCCGDLLECCADDDADLAKYICENODSISKLKCCCKPILKSHCIAEVNDMPA 300
DB 265 VHTCCGDLLECCADDDADLAKYICENODSISKLKCCCKPILKSHCIAEVNDMPA 324
QY 301 DLPSLADPVESKDVCKNVAEAKDVLGMFLYEXARHPDYSVLLRLAKTYETLEKC 360
DB 325 DLPSLADPVESKDVCKNVAEAKDVLGMFLYEXARHPDYSVLLRLAKTYETLEKC 384
QY 361 CAADPHECYAKVDEFEKPLVEEPONLIKONCELEOIGEXKFNALLVRYTKKPVOST 420
DB 385 CAADPHECYAKVDEFEKPLVEEPONLIKONCELEOIGEXKFNALLVRYTKKPVOST 444
QY 421 PTLVEVSRLNGKVGSKCKKHPKARMPQCAEDYLSVNLQCLVLEKTPVSDRYTKCTES 480
DB 445 PTLVEVSRLNGKVGSKCKKHPKARMPQCAEDYLSVNLQCLVLEKTPVSDRYTKCTES 504
QY 481 LVNRRPCFSALVEDETVVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVHKKRAT 540
DB 505 LVNRRPCFSALVEDETVVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVHKKRAT 564
QY 541 KEOLKAVMDPFAAFVEKCKKADKKECTFAEGRKKLVAASQAALGL 585
DB 565 KEOLKAVMDPFAAFVEKCKKADKKECTFAEGRKKLVAASQAALGL 609

RESULT 10
US-08-897-956A-2
Sequence 2, Application US/08897956A
Patent No. 6423512
GENERAL INFORMATION:
APPLICANT: Mary Ellen Digan
APPLICANT: Philip Lake
APPLICANT: Hermann Gram
TITLE OF INVENTION: Fusion Polypeptides
FILE REFERENCE: 600-7244/CPA
CURRENT FILING DATE: 1997-07-21
CURRENT APPLICATION NUMBER: US/08/897, 956A
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 609
TYPE: PRT
ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match 99.9%; Score 3099; DB 4; Length 609;
Best Local Similarity 99.8%; Pred. No. 1,7e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDGEEENFKALVLIAPAOYLQOCPEFDHVKLVNEVEFAKTCVADSAE 60
DB 25 DAHSEVAHFRKDGEEENFKALVLIAPAOYLQOCPEFDHVKLVNEVEFAKTCVADSAE 84
QY 61 NCKSLHTLFGDKICTYATLRETYGEMADCCAKOEPERNECFLQHKDNDNPLPLVPREV 120
DB 85 NCKSLHTLFGDKICTYATLRETYGEMADCCAKOEPERNECFLQHKDNDNPLPLVPREV 144
QY 121 DVNCTAHNDNEFTLKKYLYEIAARRHYFYAPELLFFAKRYKKAFTCCOAAKACLLP 180
DB 145 DVNCTAHNDNEFTLKKYLYEIAARRHYFYAPELLFFAKRYKKAFTCCOAAKACLLP 204
QY 181 KLDELREGKASSAKORLKCASLOKFGERAFAKANAVALRSORPKAEFAEYSKLVTDLT 240
DB 205 KLDELREGKASSAKORLKCASLOKFGERAFAKANAVALRSORPKAEFAEYSKLVTDLT 264
QY 241 VHTCCGDLLECCADDDADLAKYICENODSISKLKCCCKPILKSHCIAEVNDMPA 300
DB 265 VHTCCGDLLECCADDDADLAKYICENODSISKLKCCCKPILKSHCIAEVNDMPA 324
QY 301 DLPSLADPVESKDVCKNVAEAKDVLGMFLYEXARHPDYSVLLRLAKTYETLEKC 360
DB 325 DLPSLADPVESKDVCKNVAEAKDVLGMFLYEXARHPDYSVLLRLAKTYETLEKC 384
QY 361 CAADPHECYAKVDEFEKPLVEEPONLIKONCELEOIGEXKFNALLVRYTKKPVOST 420
DB 385 CAADPHECYAKVDEFEKPLVEEPONLIKONCELEOIGEXKFNALLVRYTKKPVOST 444
QY 421 PTLVEVSRLNGKVGSKCKKHPKARMPQCAEDYLSVNLQCLVLEKTPVSDRYTKCTES 480
DB 445 PTLVEVSRLNGKVGSKCKKHPKARMPQCAEDYLSVNLQCLVLEKTPVSDRYTKCTES 504
QY 481 LVNRRPCFSALVEDETVVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVHKKRAT 540
DB 505 LVNRRPCFSALVEDETVVPKEFNAETFTPHADICTLSEKEROIKKOTALVELVHKKRAT 564
QY 541 KEOLKAVMDPFAAFVEKCKKADKKECTFAEGRKKLVAASQAALGL 585
DB 565 KEOLKAVMDPFAAFVEKCKKADKKECTFAEGRKKLVAASQAALGL 609

RESULT 11
PCT-US95-04075-3
Sequence 3, Application PC/TUS9504075
GENERAL INFORMATION:
APPLICANT: AMGEN INC.
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Center, Patent Operations/RRC
STREET: 1840 Dehavenland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: U.S.
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04075
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids

TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US95-04075-3

Query Match 99.9%; Score 3099; DB 5; Length 609;
Best Local Similarity 99.8%; Pred. No. 1.7e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 DAHSEVAHRRKDLGEENFKALVLTAFAYLQCCPFEDHVKLVNVEFEAKTCVADESAE 60
DB 25 DAHSEVAHRRKDLGEENFKALVLTAFAYLQCCPFEDHVKLVNVEFEAKTCVADESAE 84
OY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKOPEERNECFLOKHDNPMLPRLVREY 120
DB 85 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKOPEERNECFLOKHDNPMLPRLVREY 144
OY 121 DVMCTAFHDNETFLKKTLYETIARRHPYFYAPELLFFAKRYKAFTTECCQAADKAACLLP 180
DB 145 DVMCTAFHDNETFLKKTLYETIARRHPYFYAPELLFFAKRYKAFTTECCQAADKAACLLP 204
OY 181 KLDELRDGKASSAKORLKASLQKFGERAFAKAMAVARLSORFPKAEFAEVSCLVTDLT 240
DB 205 KLDELRDGKASSAKORLKASLQKFGERAFAKAMAVARLSORFPKAEFAEVSCLVTDLT 264
OY 241 VTECCGHDLLDECADRDADLAKYICENODSISKLKCECEKPLKSHCIAEVENDEMPA 300
DB 265 VTECCGHDLLDECADRDADLAKYICENODSISKLKCECEKPLKSHCIAEVENDEMPA 324
OY 301 DLPSTLADFEVSKDCKNYAEKADVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
DB 325 DLPSTLADFEVSKDCKNYAEKADVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 384
OY 361 CAADPHHCYAKVFEDEFPVLEEPONLIKONCELEKQGEYKFNALLVRYTKKPYOYST 420
DB 385 CAADPHHCYAKVFEDEFPVLEEPONLIKONCELEKQGEYKFNALLVRYTKKPYOYST 444
OY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKTPTVSDRVTCCTES 480
DB 445 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKTPTVSDRVTCCTES 504
OY 481 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKOTALVELVKKRPAT 540
DB 505 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKOTALVELVKKRPAT 564
OY 541 KEQLAVMDDFAFYKCCCKADDKETCFABEGKRLVAASQAALGL 585
DB 565 KEQLAVMDDFAFYKCCCKADDKETCFABEGKRLVAASQAALGL 609
```

RESULT 12
US-08-897-956A-3
Sequence 3, Application US/08897956A
Patent No. 6423512
GENERAL INFORMATION:
APPLICANT: Mary Ellen Digan
APPLICANT: Philip Lake
APPLICANT: Hermann Gram
TITLE OF INVENTION: Fusion Polypeptides
FILE REFERENCE: 600-7244/GPA
CURRENT APPLICATION NUMBER: US/08/897, 956A
CURRENT FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/022,689
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 978
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Fusion polypeptide

US-08-897-956A-3

Query Match 99.7%; Score 3095; DB 4; Length 978;
Best Local Similarity 99.8%; Pred. No. 8.1e-286;
Matches 583; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 DAHSEVAHRRKDLGEENFKALVLTAFAYLQCCPFEDHVKLVNVEFEAKTCVADESAE 60
DB 212 DAHSEVAHRRKDLGEENFKALVLTAFAYLQCCPFEDHVKLVNVEFEAKTCVADESAE 271
OY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKOPEERNECFLOKHDNPMLPRLVREY 120
DB 272 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKOPEERNECFLOKHDNPMLPRLVREY 331
OY 121 DVMCTAFHDNETFLKKTLYETIARRHPYFYAPELLFFAKRYKAFTTECCQAADKAACLLP 180
DB 332 DVMCTAFHDNETFLKKTLYETIARRHPYFYAPELLFFAKRYKAFTTECCQAADKAACLLP 391
OY 181 KLDELRDGKASSAKORLKASLQKFGERAFAKAMAVARLSORFPKAEFAEVSCLVTDLT 240
DB 392 KLDELRDGKASSAKORLKASLQKFGERAFAKAMAVARLSORFPKAEFAEVSCLVTDLT 451
OY 241 VTECCGHDLLDECADRDADLAKYICENODSISKLKCECEKPLKSHCIAEVENDEMPA 300
DB 452 VTECCGHDLLDECADRDADLAKYICENODSISKLKCECEKPLKSHCIAEVENDEMPA 511
OY 301 DLPSTLADFEVSKDCKNYAEKADVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
DB 512 DLPSTLADFEVSKDCKNYAEKADVFLGMFLYEVARRHPDYSVLLRLAKTYETTLK 571
OY 361 CAADPHHCYAKVFEDEFPVLEEPONLIKONCELEKQGEYKFNALLVRYTKKPYOYST 420
DB 572 CAADPHHCYAKVFEDEFPVLEEPONLIKONCELEKQGEYKFNALLVRYTKKPYOYST 631
OY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKTPTVSDRVTCCTES 480
DB 632 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHKTPTVSDRVTCCTES 691
OY 481 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKOTALVELVKKRPAT 540
DB 692 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKOTALVELVKKRPAT 751
OY 541 KEQLAVMDDFAFYKCCCKADDKETCFABEGKRLVAASQAALG 584
DB 752 KEQLAVMDDFAFYKCCCKADDKETCFABEGKRLVAASQAALG 795
```

RESULT 13
US-08-448-196A-3
Sequence 3, Application US/08448196A
Patent No. 5780594
GENERAL INFORMATION:
APPLICANT: CARTER, DANIEL C.
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
TITLE OF INVENTION: RELATED PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: NASA
STREET: MARSHALL SPACE FLIGHT CENTER
CITY: HUNTSVILLE
STATE: ALABAMA
COUNTRY: USA
ZIP: 35812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
 NAME: BROAD JR., ROBERT L.
 REGISTRATION NUMBER: 18,757
 REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 205-544-0021
 TELEFAX: 205-544-0258
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 585 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 US-08-448-196A-3

Query Match 99.7%; Score 3093; DB 1; Length 585;
 Best Local Similarity 99.7%; Pred. No. 5,9e-286;

Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAKSEVAHRFKDLGEBNFKALVLAFAQYLQCCPEDHVKLVNEYTERAKTCVADESA 60
 DB 1 DAKSEVAHRFKDLGEBNFKALVLAFAQYLQCCPEDHVKLVNEYTERAKTCVADESA 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDDNPLRLVREY 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDDNPLRLVREY 120
 QY 121 DVNCTAFHNEETFLKYLVEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACCLP 180
 DB 121 DVNCTAFHNEETFLKYLVEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACCLP 180
 QY 122 DVNCTAFHNEETFLKYLVEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACCLP 180
 DB 122 DVNCTAFHNEETFLKYLVEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACCLP 180
 QY 181 KIDELDEGKASSAKORLKASLOKGERAFKAMAVARLSQRPKAEFAVSKLVDTLK 240
 DB 181 KIDELDEGKASSAKORLKASLOKGERAFKAMAVARLSQRPKAEFAVSKLVDTLK 240
 QY 181 KIDELDEGKASSAKORLKASLOKGERAFKAMAVARLSQRPKAEFAVSKLVDTLK 240
 DB 181 KIDELDEGKASSAKORLKASLOKGERAFKAMAVARLSQRPKAEFAVSKLVDTLK 240
 QY 241 VHTCCGHDLLECCADRADLAKYICENODISSKLECCCKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCGHDLLECCADRADLAKYICENODISSKLECCCKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPISLADFEVSKDVCNKVAEAKDVFGLMFLEYARRHPDYSVLLRLAKYETLEK 360
 DB 301 DLPISLADFEVSKDVCNKVAEAKDVFGLMFLEYARRHPDYSVLLRLAKYETLEK 360
 QY 361 CAADHDECYAKVFEDEKPLVEBPONLIKONCELFQLGKYEKFNMLLVYTKKVPVST 420
 DB 361 CAADHDECYAKVFEDEKPLVEBPONLIKONCELFQLGKYEKFNMLLVYTKKVPVST 420
 QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
 DB 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
 QY 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKPKAT 540
 DB 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKPKAT 540
 QY 541 KEOLKAVMDFAAFVEKCKADKKEKCFABEGKKLVAAASQALGL 585
 DB 541 KEOLKAVMDFAAFVEKCKADKKEKCFABEGKKLVAAASQALGL 585

RESULT 14
 US-08-984-176-1
 Sequence 1, Application US/08984176

Patent No. 5948609
 GENERAL INFORMATION:
 APPLICANT: CARTER, DANIEL C
 APPLICANT: HO, JOSEPH X
 APPLICANT: RUKER, FLORIAN
 TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
 TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

FILE REFERENCE: 08/984,176
 CURRENT APPLICATION NUMBER: US/08/984,176
 CURRENT FILING DATE: 1997-12-03
 NUMBER OF SEQ ID NOS: 1
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1
 LENGTH: 585
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-08-984-176-1

Query Match 99.7%; Score 3093; DB 2; Length 585;
 Best Local Similarity 99.7%; Pred. No. 5,9e-286;

Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAKSEVAHRFKDLGEBNFKALVLAFAQYLQCCPEDHVKLVNEYTERAKTCVADESA 60
 DB 1 DAKSEVAHRFKDLGEBNFKALVLAFAQYLQCCPEDHVKLVNEYTERAKTCVADESA 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDDNPLRLVREY 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDDNPLRLVREY 120
 QY 121 DVNCTAFHNEETFLKYLVEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACCLP 180
 DB 121 DVNCTAFHNEETFLKYLVEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACCLP 180
 QY 122 DVNCTAFHNEETFLKYLVEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACCLP 180
 DB 122 DVNCTAFHNEETFLKYLVEIARRHPYFAPPELLFFAKRYKAFTGCCAADAACCLP 180
 QY 181 KIDELDEGKASSAKORLKASLOKGERAFKAMAVARLSQRPKAEFAVSKLVDTLK 240
 DB 181 KIDELDEGKASSAKORLKASLOKGERAFKAMAVARLSQRPKAEFAVSKLVDTLK 240
 QY 241 VHTCCGHDLLECCADRADLAKYICENODISSKLECCCKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCGHDLLECCADRADLAKYICENODISSKLECCCKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPISLADFEVSKDVCNKVAEAKDVFGLMFLEYARRHPDYSVLLRLAKYETLEK 360
 DB 301 DLPISLADFEVSKDVCNKVAEAKDVFGLMFLEYARRHPDYSVLLRLAKYETLEK 360
 QY 361 CAADHDECYAKVFEDEKPLVEBPONLIKONCELFQLGKYEKFNMLLVYTKKVPVST 420
 DB 361 CAADHDECYAKVFEDEKPLVEBPONLIKONCELFQLGKYEKFNMLLVYTKKVPVST 420
 QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
 DB 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
 QY 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKPKAT 540
 DB 481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVKKPKAT 540
 QY 541 KEOLKAVMDFAAFVEKCKADKKEKCFABEGKKLVAAASQALGL 585
 DB 541 KEOLKAVMDFAAFVEKCKADKKEKCFABEGKKLVAAASQALGL 585

RESULT 15
 US-08-448-196A-5
 Sequence 5, Application US/08448196A

Patent No. 5780594
 GENERAL INFORMATION:
 APPLICANT: CARTER, DANIEL C.
 TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
 TITLE OF INVENTION: CONTRAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NASA
 STREET: MARSHALL SPACE FLIGHT CENTER
 CITY: HUNTSVILLE
 STATE: ALABAMA
 COUNTRY: USA
 ZIP: 35812

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 22, 2003, 11:44:59 ; Search time 55 Seconds

(without alignments)
1263.172 Million cell updates/sec

Title: US-09-833-118-18

Perfect score: 3103

Sequence: 1 DAHKEVAHRFKDLGPNFK.....TCFAEGCKLVASQAALCL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

1:	/cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
2:	/cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
6:	/cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
10:	/cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep1.*
11:	/cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep2.*
12:	/cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
13:	/cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
15:	/cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	585	11	US-09-929-552-2
2	3103	100.0	585	12	US-09-932-613-445
3	3103	100.0	585	12	US-09-984-010-26
4	3103	100.0	585	12	US-09-833-041-18
5	3103	100.0	585	15	US-10-153-064-5
6	3103	100.0	604	12	US-09-984-010-7
7	3103	100.0	609	12	US-09-919-039-370
8	3103	100.0	609	11	US-10-153-064-7
9	3103	100.0	610	11	US-09-984-186-2
10	3103	100.0	610	15	US-10-237-667-2
11	3103	100.0	610	15	US-10-237-708-2
12	3103	100.0	610	15	US-10-237-866-2
13	3103	100.0	610	15	US-10-237-871-2
14	3103	100.0	610	15	US-10-237-624-2
15	3103	100.0	651	15	US-10-153-064-133
16	3103	100.0	652	15	US-10-153-064-132

17	3103	100.0	653	15	US-10-153-064-131	Sequence 131, App
18	3103	100.0	656	15	US-10-153-064-130	Sequence 130, App
19	3103	100.0	676	15	US-10-153-064-127	Sequence 127, App
20	3103	100.0	676	15	US-10-153-064-129	Sequence 129, App
21	3103	100.0	677	15	US-10-153-064-125	Sequence 125, App
22	3103	100.0	680	15	US-10-153-064-123	Sequence 123, App
23	3103	100.0	787	11	US-09-984-186-16	Sequence 16, App1
24	3103	100.0	787	15	US-10-237-667-16	Sequence 16, App1
25	3103	100.0	787	15	US-10-237-708-16	Sequence 16, App1
26	3103	100.0	787	15	US-10-237-866-16	Sequence 16, App1
27	3103	100.0	787	15	US-10-237-871-16	Sequence 16, App1
28	3103	100.0	787	15	US-10-237-624-16	Sequence 16, App1
29	3103	100.0	788	15	US-10-073-118-26	Sequence 26, App1
30	3092.5	99.7	652	15	US-10-153-064-96	Sequence 96, App1
31	3092.5	99.7	652	15	US-10-153-064-99	Sequence 99, App1
32	3092.5	99.7	652	15	US-10-153-064-105	Sequence 105, App
33	3092.5	99.7	660	15	US-10-153-064-90	Sequence 90, App1
34	3092.5	99.7	660	15	US-10-153-064-93	Sequence 93, App1
35	3092.5	99.7	676	15	US-10-153-064-95	Sequence 95, App1
36	3092.5	99.7	676	15	US-10-153-064-98	Sequence 98, App1
37	3092.5	99.7	676	15	US-10-153-064-104	Sequence 104, App
38	3092.5	99.7	684	15	US-10-153-064-92	Sequence 92, App1
39	3092.5	99.7	1184	15	US-10-153-064-89	Sequence 89, App1
40	3086.5	99.5	668	15	US-10-153-064-102	Sequence 102, App
41	3086.5	99.5	692	15	US-10-153-064-101	Sequence 101, App
42	2450.5	79.0	607	15	US-10-057-789-45	Sequence 45, App1
43	2450.5	79.0	607	15	US-10-212-628-45	Sequence 45, App1
44	2426	78.2	608	15	US-10-165-603-24	Sequence 24, App1
45	2426	78.2	608	15	US-10-165-603-25	Sequence 25, App1

ALIGNMENTS

RESULT 1
US-09-929-552-2
Sequence 2, Application US/09929552
Patent No. US20020123080A1

GENERAL INFORMATION:
APPLICANT: Sonnenschein, Carlos
Soto, Ana M.
TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/929,552
FILING DATE: 14-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/769,746
FILING DATE: 19-DEC-1996

ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: MBRI-02584
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.8e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 DAHSEVAHRRKRDGEENFKALVLIAPAYLQOCPFEEDHVKLVNVEVFAATCVADEAE 60
DB 1 DAHSEVAHRRKRDGEENFKALVLIAPAYLQOCPFEEDHVKLVNVEVFAATCVADEAE 60
QY 61 NCDSLHFLFDKICTVATLTRETYGEMADCCAKOEPENECFLOHKDNPMLPRLVREY 120
DB 61 NCDSLHFLFDKICTVATLTRETYGEMADCCAKOEPENECFLOHKDNPMLPRLVREY 120
QY 121 DVMCATFHDNETFLKTYELIARHPYFVAPPELLFFAKRYKAATFECQADRAACLLP 180
DB 121 DVMCATFHDNETFLKTYELIARHPYFVAPPELLFFAKRYKAATFECQADRAACLLP 180
QY 121 DVMCATFHDNETFLKTYELIARHPYFVAPPELLFFAKRYKAATFECQADRAACLLP 180
DB 121 DVMCATFHDNETFLKTYELIARHPYFVAPPELLFFAKRYKAATFECQADRAACLLP 180
QY 181 KLDELRDGKASAKORLKASLQKFGERRAKAMAVARLSQFPRAEFAEYSKLVTDLTK 240
DB 181 KLDELRDGKASAKORLKASLQKFGERRAKAMAVARLSQFPRAEFAEYSKLVTDLTK 240
QY 241 VHTCCGHDLLCADDRAADLAKYICENODSISKKECEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLCADDRAADLAKYICENODSISKKECEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFEVSKDCKYAKAKDVFGLMFLYETARRHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLAADFEVSKDCKYAKAKDVFGLMFLYETARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHBCYAKVDEFEPLVEEPONLIKONCELEFQGEYKFOALLVRYTKKVPQYST 420
DB 361 CAADPHBCYAKVDEFEPLVEEPONLIKONCELEFQGEYKFOALLVRYTKKVPQYST 420
QY 421 PTLVEVSNNLKGVSCKCKHBEAKRMPCAEDYLSVNLQCVLHKTPTVSRYVKCTES 480
DB 421 PTLVEVSNNLKGVSCKCKHBEAKRMPCAEDYLSVNLQCVLHKTPTVSRYVKCTES 480
QY 481 LVNRRPCSALEVEDETYPKFEFNAETFTFHADICTLSKEKROIKQOTALVELYKHPKAT 540
DB 481 LVNRRPCSALEVEDETYPKFEFNAETFTFHADICTLSKEKROIKQOTALVELYKHPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCFAEEGKKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDKETCFAEEGKKLVAAASQAALGL 585

RESULT 2
US-09-932-613-445
; Sequence 445, Application US/09932613
; Publication No. US20030091565A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Beltzer, James P.
; APPLICANT: Pottery, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: DXX-025.1 PCT: DXX-025.1 US
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-613-445
Query Match 100.0%; Score 3103; DB 12; Length 585;

```

```

Best Local Similarity 100.0%; Pred. No. 7.8e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRKRDGEENFKALVLIAPAYLQOCPFEEDHVKLVNVEVFAATCVADEAE 60
DB 1 DAHSEVAHRRKRDGEENFKALVLIAPAYLQOCPFEEDHVKLVNVEVFAATCVADEAE 60
QY 61 NCDSLHFLFDKICTVATLTRETYGEMADCCAKOEPENECFLOHKDNPMLPRLVREY 120
DB 61 NCDSLHFLFDKICTVATLTRETYGEMADCCAKOEPENECFLOHKDNPMLPRLVREY 120
QY 121 DVMCATFHDNETFLKTYELIARHPYFVAPPELLFFAKRYKAATFECQADRAACLLP 180
DB 121 DVMCATFHDNETFLKTYELIARHPYFVAPPELLFFAKRYKAATFECQADRAACLLP 180
QY 121 DVMCATFHDNETFLKTYELIARHPYFVAPPELLFFAKRYKAATFECQADRAACLLP 180
DB 121 DVMCATFHDNETFLKTYELIARHPYFVAPPELLFFAKRYKAATFECQADRAACLLP 180
QY 181 KLDELRDGKASAKORLKASLQKFGERRAKAMAVARLSQFPRAEFAEYSKLVTDLTK 240
DB 181 KLDELRDGKASAKORLKASLQKFGERRAKAMAVARLSQFPRAEFAEYSKLVTDLTK 240
QY 241 VHTCCGHDLLCADDRAADLAKYICENODSISKKECEKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLCADDRAADLAKYICENODSISKKECEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFEVSKDCKYAKAKDVFGLMFLYETARRHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLAADFEVSKDCKYAKAKDVFGLMFLYETARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHBCYAKVDEFEPLVEEPONLIKONCELEFQGEYKFOALLVRYTKKVPQYST 420
DB 361 CAADPHBCYAKVDEFEPLVEEPONLIKONCELEFQGEYKFOALLVRYTKKVPQYST 420
QY 421 PTLVEVSNNLKGVSCKCKHBEAKRMPCAEDYLSVNLQCVLHKTPTVSRYVKCTES 480
DB 421 PTLVEVSNNLKGVSCKCKHBEAKRMPCAEDYLSVNLQCVLHKTPTVSRYVKCTES 480
QY 481 LVNRRPCSALEVEDETYPKFEFNAETFTFHADICTLSKEKROIKQOTALVELYKHPKAT 540
DB 481 LVNRRPCSALEVEDETYPKFEFNAETFTFHADICTLSKEKROIKQOTALVELYKHPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCFAEEGKKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDKETCFAEEGKKLVAAASQAALGL 585

RESULT 3
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESS: FINNEGAN, HENDERSON, PARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984, 010
; FILING DATE: 21-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091, 873
; FILING DATE: 25-JUN-1998
; APPLICATION NUMBER: PCT/GB96/03164
; FILING DATE: 19-DEC-1996

```

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.8e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 DAHSEVAHREFKDLGEEFNKALVLIAPQYLQCCPEPDHVKLVNEVTEFAKTCVADSAE 60
DB 1 DAHSEVAHREFKDLGEEFNKALVLIAPQYLQCCPEPDHVKLVNEVTEFAKTCVADSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCFLOHKDNDNPLRLVREY 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCFLOHKDNDNPLRLVREY 120
QY 121 DVNCTAFHNEETFLKKYLYEIAARRHPFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
DB 121 DVNCTAFHNEETFLKKYLYEIAARRHPFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
QY 181 KIDELRDEGKASSAKORLKCASLOKGERAFKAMAAVARLSQRPKAEFAVSKLYVDLTK 240
DB 181 KIDELRDEGKASSAKORLKCASLOKGERAFKAMAAVARLSQRPKAEFAVSKLYVDLTK 240
QY 241 VHTTECGHGLLECGADBRADLAKYICENODSISKLECKCKPILKSHCIAEVNDMPA 300
DB 241 VHTTECGHGLLECGADBRADLAKYICENODSISKLECKCKPILKSHCIAEVNDMPA 300
QY 301 DLPSLAADFVESKDVCNKVAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYETLLEKC 360
DB 301 DLPSLAADFVESKDVCNKVAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYETLLEKC 360
QY 361 CAADPHECYAKVFDEKPLVEEPONLIKONCELFEOLGEXKFNOMALLVYTRKVPVST 420
DB 361 CAADPHECYAKVFDEKPLVEEPONLIKONCELFEOLGEXKFNOMALLVYTRKVPVST 420
QY 421 PTLVEVSRLNGKSGKCKRHPKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEVSRLNGKSGKCKRHPKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPCFSALVEDETVVPKEFNAETFTFHADICTISEKERQIKKOTALVELVKKHKPAT 540
DB 481 LVNRRPCFSALVEDETVVPKEFNAETFTFHADICTISEKERQIKKOTALVELVKKHKPAT 540
QY 541 KEOLKAVMDFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL 585
DB 541 KEOLKAVMDFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL 585
```

RESULT 4
US-09-833-041-18
Sequence 18, Application US/09833041
Publication No. US20030125247A1
GENERAL INFORMATION:
APPLICANT: Rosen, Craig A.
APPLICANT: Haseltine, William A.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF545
CURRENT APPLICATION NUMBER: US/09/833,041
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384

PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ. ID NOS: 79
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 18
LENGTH: 585
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-833-041-18

Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.8e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 DAHSEVAHREFKDLGEEFNKALVLIAPQYLQCCPEPDHVKLVNEVTEFAKTCVADSAE 60
DB 1 DAHSEVAHREFKDLGEEFNKALVLIAPQYLQCCPEPDHVKLVNEVTEFAKTCVADSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCFLOHKDNDNPLRLVREY 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCFLOHKDNDNPLRLVREY 120
QY 121 DVNCTAFHNEETFLKKYLYEIAARRHPFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
DB 121 DVNCTAFHNEETFLKKYLYEIAARRHPFYAPPELLFFAKRYKAAFTCCOAAADKAACLLP 180
QY 181 KIDELRDEGKASSAKORLKCASLOKGERAFKAMAAVARLSQRPKAEFAVSKLYVDLTK 240
DB 181 KIDELRDEGKASSAKORLKCASLOKGERAFKAMAAVARLSQRPKAEFAVSKLYVDLTK 240
QY 241 VHTTECGHGLLECGADBRADLAKYICENODSISKLECKCKPILKSHCIAEVNDMPA 300
DB 241 VHTTECGHGLLECGADBRADLAKYICENODSISKLECKCKPILKSHCIAEVNDMPA 300
QY 301 DLPSLAADFVESKDVCNKVAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYETLLEKC 360
DB 301 DLPSLAADFVESKDVCNKVAEAKDVFGLMFLYEYARRHPDYSVLLRLAKTYETLLEKC 360
QY 361 CAADPHECYAKVFDEKPLVEEPONLIKONCELFEOLGEXKFNOMALLVYTRKVPVST 420
DB 361 CAADPHECYAKVFDEKPLVEEPONLIKONCELFEOLGEXKFNOMALLVYTRKVPVST 420
QY 421 PTLVEVSRLNGKSGKCKRHPKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEVSRLNGKSGKCKRHPKRMPCADYLSVNLQCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPCFSALVEDETVVPKEFNAETFTFHADICTISEKERQIKKOTALVELVKKHKPAT 540
DB 481 LVNRRPCFSALVEDETVVPKEFNAETFTFHADICTISEKERQIKKOTALVELVKKHKPAT 540
QY 541 KEOLKAVMDFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL 585
DB 541 KEOLKAVMDFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL 585
```

RESULT 5
US-10-153-064-5
Sequence 5, Application US/10153064
Publication No. US20020142814A1
GENERAL INFORMATION:
APPLICANT: Bell et al.
TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
FILE REFERENCE: PF556
CURRENT APPLICATION NUMBER: US/10/153,064
CURRENT FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: 60/293,212
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ. ID NOS: 137
SOFTWARE: Patent In version 3.1
SEQ ID NO 5
LENGTH: 585
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 3103; DB 15; Length 585;
 Best Local Similarity 100.0%; Pred. No. 7, 8e-271;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 DAHSEVAHRRKIDGEGENFKALVLIAPAOYLQCCPFEDHVKLVNEVEFAKTCVADESA 60
DB 1 DAHSEVAHRRKIDGEGENFKALVLIAPAOYLQCCPFEDHVKLVNEVEFAKTCVADESA 60
OY 61 NCDSLHTLFGDKLCTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVRE 120
DB 61 NCDSLHTLFGDKLCTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVRE 120
OY 121 DVMTAFHDNEETFLKRYLFIARRHPFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
DB 121 DVMTAFHDNEETFLKRYLFIARRHPFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
OY 181 KLDELREGKASSAKORLKASLQKFGFRAKANAVARLSORFPKAEVSKLVTDLT 240
DB 181 KLDELREGKASSAKORLKASLQKFGFRAKANAVARLSORFPKAEVSKLVTDLT 240
OY 241 VHTECCHDILLECADDRLADLAICYCENODSISSKLKECEKPLEKSHCIAVENDEMPA 300
DB 241 VHTECCHDILLECADDRLADLAICYCENODSISSKLKECEKPLEKSHCIAVENDEMPA 300
OY 301 DLPSLAADFESKDVCKNYAKADVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLAADFESKDVCKNYAKADVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360
OY 361 CAADPHCEYAKVDEFPPLVEEPONLIKONCELFEOUGEYKFPONALLVRYTKKVPVST 420
DB 361 CAADPHCEYAKVDEFPPLVEEPONLIKONCELFEOUGEYKFPONALLVRYTKKVPVST 420
OY 421 PTIVEVSNNLKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSDRYTKCTES 480
DB 421 PTIVEVSNNLKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSDRYTKCTES 480
OY 481 LVNRRPCSALEVEDETVPKPEFNAETFTFHADICTLSKEKQIKQOTALVELVKKPKAT 540
DB 481 LVNRRPCSALEVEDETVPKPEFNAETFTFHADICTLSKEKQIKQOTALVELVKKPKAT 540
OY 541 KEQLKAVMDDEFAAFVEKCKADDKETCFABEGKRLVAASQAALGL 585
DB 541 KEQLKAVMDDEFAAFVEKCKADDKETCFABEGKRLVAASQAALGL 585

```

RESULT 6

US-09-984-010-7
 ; Sequence 7, Application US/09984010
 ; Publication No. US20030104578A1

GENERAL INFORMATION:

APPLICANT: Ballance, David James
 TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
 AND SERUM ALBUMIN

NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:

ADDRESSSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
 STREET: 1300 I Street, NW
 CITY: Washington
 STATE: DC

COUNTRY: USA
 ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/984, 010

FILING DATE: 21-May-2002
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/091, 873
 FILING DATE: 25-JUN-1998

APPLICATION NUMBER: PCT/G996/03164
 FILING DATE: 19-DEC-1996
 INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 604 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHEetical: NO

AMTI-SENSE: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-984-010-7
 Query Match 100.0%; Score 3103; DB 12; Length 604;
 Best Local Similarity 100.0%; Pred. No. 8, 2e-271;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 DAHSEVAHRRKIDGEGENFKALVLIAPAOYLQCCPFEDHVKLVNEVEFAKTCVADESA 60
DB 20 DAHSEVAHRRKIDGEGENFKALVLIAPAOYLQCCPFEDHVKLVNEVEFAKTCVADESA 79
OY 61 NCDSLHTLFGDKLCTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVRE 120
DB 80 NCDSLHTLFGDKLCTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVRE 139
OY 121 DVMTAFHDNEETFLKRYLFIARRHPFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 180
DB 140 DVMTAFHDNEETFLKRYLFIARRHPFYAPPELLFFAKRYKAAFTTECCQAADRAACILP 199
OY 181 KLDELREGKASSAKORLKASLQKFGFRAKANAVARLSORFPKAEVSKLVTDLT 240
DB 200 KLDELREGKASSAKORLKASLQKFGFRAKANAVARLSORFPKAEVSKLVTDLT 259
OY 241 VHTECCHDILLECADDRLADLAICYCENODSISSKLKECEKPLEKSHCIAVENDEMPA 300
DB 260 VHTECCHDILLECADDRLADLAICYCENODSISSKLKECEKPLEKSHCIAVENDEMPA 319
OY 301 DLPSLAADFESKDVCKNYAKADVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360
DB 320 DLPSLAADFESKDVCKNYAKADVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 379
OY 361 CAADPHCEYAKVDEFPPLVEEPONLIKONCELFEOUGEYKFPONALLVRYTKKVPVST 420
DB 380 CAADPHCEYAKVDEFPPLVEEPONLIKONCELFEOUGEYKFPONALLVRYTKKVPVST 439
OY 421 PTIVEVSNNLKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSDRYTKCTES 480
DB 440 PTIVEVSNNLKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPVSDRYTKCTES 499
OY 481 LVNRRPCSALEVEDETVPKPEFNAETFTFHADICTLSKEKQIKQOTALVELVKKPKAT 540
DB 500 LVNRRPCSALEVEDETVPKPEFNAETFTFHADICTLSKEKQIKQOTALVELVKKPKAT 559
OY 541 KEQLKAVMDDEFAAFVEKCKADDKETCFABEGKRLVAASQAALGL 585
DB 560 KEQLKAVMDDEFAAFVEKCKADDKETCFABEGKRLVAASQAALGL 604

```

RESULT 7

US-09-919-039-370
 ; Sequence 370, Application US/09919039
 ; Publication No. US20030108871A1

GENERAL INFORMATION:

APPLICANT: Kaser, Matthew R.

TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES

FILE REFERENCE: PA-0035 US

CURRENT APPLICATION NUMBER: US/09/919, 039
 CURRENT FILING DATE: 2002-09-09
 PRIOR APPLICATION NUMBER: 60/222, 113
 PRIOR FILING DATE: 2000-07-28
 NUMBER OF SEQ ID NOS: 401
 SOFTWARE: PERL Program

```

; SEQ ID NO 370
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 088957CD1
US-09-919-039-370

```

```

Query Match      100.0%; Score 3103; DB 12; Length 609;
Best Local Similarity 100.0%; Pred. No. 8.3e-271;
Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPREDHYKLVNEVTERAKTVADESAS 60
DB 25 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPREDHYKLVNEVTERAKTVADESAS 84
QY 61 NCDSKSLHTLFQDKLCTVATLRETYGEMADCCAKQEBERNECFLOHDDNDPNLPRLVREY 120
DB 85 NCDSKSLHTLFQDKLCTVATLRETYGEMADCCAKQEBERNECFLOHDDNDPNLPRLVREY 144
QY 121 DWACSTAFHNEEFELKKYIETARRHPYAPPLLEFFARKYKAFTCCOADAADKACLP 180
DB 145 DWACSTAFHNEEFELKKYIETARRHPYAPPLLEFFARKYKAFTCCOADAADKACLP 204
QY 181 KIDELDEBKASSAKORLKCASLOKFGERAFAKMAVARLSORPFAKFAEVSCLVTDLR 240
DB 205 KIDELDEBKASSAKORLKCASLOKFGERAFAKMAVARLSORPFAKFAEVSCLVTDLR 264
QY 241 VHTECCHGDLLECCADRADLAKYICENODSISSKLECCCKPFLKSHCIAEYENDMPA 300
DB 265 VHTECCHGDLLECCADRADLAKYICENODSISSKLECCCKPFLKSHCIAEYENDMPA 324
QY 301 DLPSTLADVESDVCNKNAEAKDVFLGMFLIYARHPDYSVLLRLAKYETTLK 360
DB 325 DLPSTLADVESDVCNKNAEAKDVFLGMFLIYARHPDYSVLLRLAKYETTLK 384
QY 361 CAADHHECYAKYFDEFKPLVEBPONLIONCELFQJLGEYKQNALVRYTKKPVQVST 420
DB 385 CAADHHECYAKYFDEFKPLVEBPONLIONCELFQJLGEYKQNALVRYTKKPVQVST 444
QY 421 PTLVEYSRNLGKVGSKCKHPKAPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 445 PTLVEYSRNLGKVGSKCKHPKAPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 504
QY 481 LVNRRPCFSALVEDEYVKEFNAETFTFHADICTISEKEROIKKOTALVEIYKHPKAT 540
DB 505 LVNRRPCFSALVEDEYVKEFNAETFTFHADICTISEKEROIKKOTALVEIYKHPKAT 564
QY 541 KEOLKAVMDPFAAFVEKCKKADDKETCFABEGKRLVAASQAALGL 585
DB 565 KEOLKAVMDPFAAFVEKCKKADDKETCFABEGKRLVAASQAALGL 609

```

RESULT 8

```

US-10-153-064-7
; Sequence 7, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-7

```

```

Query Match      100.0%; Score 3103; DB 15; Length 609;
Best Local Similarity 100.0%; Pred. No. 8.3e-271;
Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPREDHYKLVNEVTERAKTVADESAS 60
DB 25 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPREDHYKLVNEVTERAKTVADESAS 84
QY 61 NCDSKSLHTLFQDKLCTVATLRETYGEMADCCAKQEBERNECFLOHDDNDPNLPRLVREY 120
DB 85 NCDSKSLHTLFQDKLCTVATLRETYGEMADCCAKQEBERNECFLOHDDNDPNLPRLVREY 144
QY 121 DWACSTAFHNEEFELKKYIETARRHPYAPPLLEFFARKYKAFTCCOADAADKACLP 180
DB 145 DWACSTAFHNEEFELKKYIETARRHPYAPPLLEFFARKYKAFTCCOADAADKACLP 204
QY 181 KIDELDEBKASSAKORLKCASLOKFGERAFAKMAVARLSORPFAKFAEVSCLVTDLR 240
DB 205 KIDELDEBKASSAKORLKCASLOKFGERAFAKMAVARLSORPFAKFAEVSCLVTDLR 264
QY 241 VHTECCHGDLLECCADRADLAKYICENODSISSKLECCCKPFLKSHCIAEYENDMPA 300
DB 265 VHTECCHGDLLECCADRADLAKYICENODSISSKLECCCKPFLKSHCIAEYENDMPA 324
QY 301 DLPSTLADVESDVCNKNAEAKDVFLGMFLIYARHPDYSVLLRLAKYETTLK 360
DB 325 DLPSTLADVESDVCNKNAEAKDVFLGMFLIYARHPDYSVLLRLAKYETTLK 384
QY 361 CAADHHECYAKYFDEFKPLVEBPONLIONCELFQJLGEYKQNALVRYTKKPVQVST 420
DB 385 CAADHHECYAKYFDEFKPLVEBPONLIONCELFQJLGEYKQNALVRYTKKPVQVST 444
QY 421 PTLVEYSRNLGKVGSKCKHPKAPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 480
DB 445 PTLVEYSRNLGKVGSKCKHPKAPCAEDYLSVNLQCVLHEKTPVSDRYTKCTES 504
QY 481 LVNRRPCFSALVEDEYVKEFNAETFTFHADICTISEKEROIKKOTALVEIYKHPKAT 540
DB 505 LVNRRPCFSALVEDEYVKEFNAETFTFHADICTISEKEROIKKOTALVEIYKHPKAT 564
QY 541 KEOLKAVMDPFAAFVEKCKKADDKETCFABEGKRLVAASQAALGL 585
DB 565 KEOLKAVMDPFAAFVEKCKKADDKETCFABEGKRLVAASQAALGL 609

```

RESULT 9

```

US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. US2002015101A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Patrice
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcoia Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match 100.0%; Score 3103; DB 11; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.3e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DAHSEVAHRRKDDGEENFKALVLIARFQYIQCCPFEDHVKLVNEVEFAKTCVADSAAE 60
25 DAHSEVAHRRKDDGEENFKALVLIARFQYIQCCPFEDHVKLVNEVEFAKTCVADSAAE 84
61 NCDSKSLHTLFGDKLCTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVREPV 120
85 NCDSKSLHTLFGDKLCTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVREPV 144
121 DVMCSTAFDNEETFLKLYEIAARHPFYFAPPELLFFAKRYKAFTTECCQAADRAACLLP 180
145 DVMCSTAFDNEETFLKLYEIAARHPFYFAPPELLFFAKRYKAFTTECCQAADRAACLLP 204
181 KLDELROEGKASSAKORLKASLOKFGBRAKAMAVARLSORFPAEAEVSKLVTDLTK 240
205 KLDELROEGKASSAKORLKASLOKFGBRAKAMAVARLSORFPAEAEVSKLVTDLTK 264
241 VHTCCGHDDLCEADDRADLAKYICENDSISIKLKECCERPLEKSHCIAEVENDENPA 300
265 VHTCCGHDDLCEADDRADLAKYICENDSISIKLKECCERPLEKSHCIAEVENDENPA 324
301 DLPSLAADFVSKOVCKNYAKADVFLGMFLYEYARRHPDVSVYLLRLATYETTLK 360
325 DLPSLAADFVSKOVCKNYAKADVFLGMFLYEYARRHPDVSVYLLRLATYETTLK 384
361 CAADPHHCYAKVDFEFLVEEPONLIKONCELFEOLGEYKFOVALVRYTKKPVYST 420
385 CAADPHHCYAKVDFEFLVEEPONLIKONCELFEOLGEYKFOVALVRYTKKPVYST 444
421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTPEVSDRVTKCTES 480
445 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTPEVSDRVTKCTES 504
481 LVNRRPCSALEVEDTYPRKFNATFTFHADICTLSKEROIKQOTLVLYLVKHKPKAT 540
505 LVNRRPCSALEVEDTYPRKFNATFTFHADICTLSKEROIKQOTLVLYLVKHKPKAT 564
541 KEOLKAVMDPFAAFYEKCCRAADKETCFPAEGSKLVAASOALGL 585
565 KEOLKAVMDPFAAFYEKCCRAADKETCFPAEGSKLVAASOALGL 609

RESULT 10
US-10-237-667-2
Sequence 2, Application US/10237667
Publication No. US20030022308A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES.
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3543
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,667
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-667-2

Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.3e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DAHSEVAHRRKDDGEENFKALVLIARFQYIQCCPFEDHVKLVNEVEFAKTCVADSAAE 60
25 DAHSEVAHRRKDDGEENFKALVLIARFQYIQCCPFEDHVKLVNEVEFAKTCVADSAAE 84
61 NCDSKSLHTLFGDKLCTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVREPV 120
85 NCDSKSLHTLFGDKLCTVAATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVREPV 144
121 DVMCSTAFDNEETFLKLYEIAARHPFYFAPPELLFFAKRYKAFTTECCQAADRAACLLP 180
145 DVMCSTAFDNEETFLKLYEIAARHPFYFAPPELLFFAKRYKAFTTECCQAADRAACLLP 204
181 KLDELROEGKASSAKORLKASLOKFGBRAKAMAVARLSORFPAEAEVSKLVTDLTK 240


```

Db      205 KIDELDEGKASSAKORLKCASLOKGERAFKAMAVARLSORPFAEVSCLVTLTK 264
QY      241 VTECHGHDLLBCADRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEYENDMPA 300
Db      265 VTECHGHDLLBCADRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEYENDMPA 324
QY      301 DLPSLAADVESKDVCKNNAEAKDVFLGMFLYEYARRHDPYSVLLRLAKYETTLK 360
Db      325 DLPSLAADVESKDVCKNNAEAKDVFLGMFLYEYARRHDPYSVLLRLAKYETTLK 384
QY      361 CAADPHCEYAKVFDEKPLVEEPONLIKONCELEFOLGEYKFNALLVRYTKYQVST 420
Db      385 CAADPHCEYAKVFDEKPLVEEPONLIKONCELEFOLGEYKFNALLVRYTKYQVST 444
QY      421 PTLVEYSRLKGVGSKCKHPEAKRMPKCAEDYLSVNLQCVLHETPVS DRVTKCTES 480
Db      445 PTLVEYSRLKGVGSKCKHPEAKRMPKCAEDYLSVNLQCVLHETPVS DRVTKCTES 504
QY      481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKPKAT 540
Db      505 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKPKAT 564
QY      541 KEQLAVMDFAAFVEKCKKADDKETCFAEESKRLVAASQAALGL 585
Db      565 KEQLAVMDFAAFVEKCKKADDKETCFAEESKRLVAASQAALGL 609

```

RESULT 11
US-10-237-708-2
Sequence 2, Application US/10237708
Publication No. US20030036170A1

GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard
Fournier, Alain
Guiton, Jean-Dominique
Jung, Gerard
Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Inc.

STREET: 500 Arcola Road, 3CA3

CITY: Collegeville

STATE: PA

COUNTRY: USA

ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: System 7.1

SOFTWARE: Word 5.1 (PatentIn)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/237, 708

FILING DATE: 10-SEP-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION NUMBER:

APPLICATION NUMBER: US/08/797, 689

FILING DATE: 31-JAN-1997

APPLICATION NUMBER: US 08/256, 927

FILING DATE: 28-JUL-1994

APPLICATION NUMBER: FR 92/01064

FILING DATE: 31-JAN-1992

APPLICATION NUMBER: PCT/FR93/00085

FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Smith Ph. D., Julie K.

REGISTRATION NUMBER: P-38, 619

REFERENCE/DOCKET NUMBER: ST92006-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (610) 454-3839

TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-708-2

Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.3e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 DAHKSFAHFPDIDGENFKALVTLFAFQYLQCCPEHDHVKLVNTEERAKVCVADSEAE 60
Db      25 DAHKSFAHFPDIDGENFKALVTLFAFQYLQCCPEHDHVKLVNTEERAKVCVADSEAE 84
QY      61 NCDKSLHTLFGDKLCTVATLRETYGMAADCCAKQOEERNECFLOHKDNPMLPRLVREPV 120
Db      85 NCDKSLHTLFGDKLCTVATLRETYGMAADCCAKQOEERNECFLOHKDNPMLPRLVREPV 144
QY      121 DVMCTAFHNEETFLKLYEYIARRHPYYPABELFFAKRYAAFTCCQAADKAACLLP 180
Db      145 DVMCTAFHNEETFLKLYEYIARRHPYYPABELFFAKRYAAFTCCQAADKAACLLP 204
QY      181 KIDELDEGKASSAKORLKCASLOKGERAFKAMAVARLSORPFAEVSCLVTLTK 240
Db      205 KIDELDEGKASSAKORLKCASLOKGERAFKAMAVARLSORPFAEVSCLVTLTK 264
QY      241 VTECHGHDLLBCADRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEYENDMPA 300
Db      265 VTECHGHDLLBCADRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEYENDMPA 324
QY      301 DLPSLAADVESKDVCKNNAEAKDVFLGMFLYEYARRHDPYSVLLRLAKYETTLK 360
Db      325 DLPSLAADVESKDVCKNNAEAKDVFLGMFLYEYARRHDPYSVLLRLAKYETTLK 384
QY      361 CAADPHCEYAKVFDEKPLVEEPONLIKONCELEFOLGEYKFNALLVRYTKYQVST 420
Db      385 CAADPHCEYAKVFDEKPLVEEPONLIKONCELEFOLGEYKFNALLVRYTKYQVST 444
QY      421 PTLVEYSRLKGVGSKCKHPEAKRMPKCAEDYLSVNLQCVLHETPVS DRVTKCTES 480
Db      445 PTLVEYSRLKGVGSKCKHPEAKRMPKCAEDYLSVNLQCVLHETPVS DRVTKCTES 504
QY      481 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKPKAT 540
Db      505 LVNRRPCFSALVEDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVKKPKAT 564
QY      541 KEQLAVMDFAAFVEKCKKADDKETCFAEESKRLVAASQAALGL 585
Db      565 KEQLAVMDFAAFVEKCKKADDKETCFAEESKRLVAASQAALGL 609

```

RESULT 12
US-10-237-866-2
Sequence 2, Application US/10237866
Publication No. US20030036171A1

GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard
Fournier, Alain
Guiton, Jean-Dominique
Jung, Gerard
Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Inc.

STREET: 500 Arcola Road, 3CA3

CITY: Collegeville

STATE: PA

COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (patentln)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-866-2

Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.3e-211;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRKDKLGEENFKALVLIARFQYIQQCPFEDHVKLVNEVTEFAKTCVADSAAE 60
DB 25 DAHSEVAHRRKDKLGEENFKALVLIARFQYIQQCPFEDHVKLVNEVTEFAKTCVADSAAE 84
QY 61 NCDSKSLHTLFGDKICTYATLTRETYGEMADCCAKOEPENECFLQHKDNPMLPRLVREPV 120
DB 85 NCDSKSLHTLFGDKICTYATLTRETYGEMADCCAKOEPENECFLQHKDNPMLPRLVREPV 144
QY 121 DVMCTAFHDNEETFLKLYEYIARRHPYFYAPPELLFPAKRYKAAFTCCQAADRAACILP 180
DB 145 DVMCTAFHDNEETFLKLYEYIARRHPYFYAPPELLFPAKRYKAAFTCCQAADRAACILP 204
QY 181 KLDELROEGKASSAKQRLKASLQKFGFRARAKANAVALSORFPAKAEAEVSKLVTDTLK 240
DB 205 KLDELROEGKASSAKQRLKASLQKFGFRARAKANAVALSORFPAKAEAEVSKLVTDTLK 264
QY 241 VHTCCGHDLLCECADDRADLAKYICENODSISSKLKECCCKPLEKSCIAEVENDEMPA 300
DB 265 VHTCCGHDLLCECADDRADLAKYICENODSISSKLKECCCKPLEKSCIAEVENDEMPA 324
QY 301 DLPSLAADVFESKDVCKNYAEAKDVFLGMLLEYAARRHDDYSVLLRLAATYETLEK 360
DB 325 DLPSLAADVFESKDVCKNYAEAKDVFLGMLLEYAARRHDDYSVLLRLAATYETLEK 384
QY 361 CAADPHHCYAKVDEFRPLVEEPONLIKONCELFEOJGEEKFQONALLVRTTKVPOYST 420
DB 385 CAADPHHCYAKVDEFRPLVEEPONLIKONCELFEOJGEEKFQONALLVRTTKVPOYST 444
QY 421 PTLVEVSNTLKVSSCKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPDSRVTKCCES 480
DB 445 PTLVEVSNTLKVSSCKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTVPDSRVTKCCES 504
QY 481 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSERQIKQOTALVELVVKHPRKAT 540

DB 505 LVNRRPCSALEVDETYVPKEFNAETFTFHADICTLSERQIKQOTALVELVVKHPRKAT 564
QY 541 KEOLKAVMDPFAAFVEKCKKADKRETCFAEGKRLVAASQAALGL 585
DB 565 KEOLKAVMDPFAAFVEKCKKADKRETCFAEGKRLVAASQAALGL 609

RESULT 13
US-10-237-871-2
Sequence 2, Application US/10237871
Publication No. US20030036172A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patricia
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcoia Road, 3043
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (patentln)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,871
FILING DATE: 10-Sep-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-871-2

Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.3e-211;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRKDKLGEENFKALVLIARFQYIQQCPFEDHVKLVNEVTEFAKTCVADSAAE 60
DB 25 DAHSEVAHRRKDKLGEENFKALVLIARFQYIQQCPFEDHVKLVNEVTEFAKTCVADSAAE 84
QY 61 NCDSKSLHTLFGDKICTYATLTRETYGEMADCCAKOEPENECFLQHKDNPMLPRLVREPV 120

```

Db      85 NCDKSLHTLFSGDKLCTVATLRETYGEMADCCAKOBERNECEFLQHKDNDPNI.PRLVREPV 144
QY      121 DWGCTAFHNEETFLKKYLYETARRRPYAPPELLFFAKRYKAAPFECOAADKAACLP 180
Db      145 DWGCTAFHNEETFLKKYLYETARRRPYAPPELLFFAKRYKAAPFECOAADKAACLP 204
QY      181 KIDELDEBKASSAKORLKASLOKFGERAFAKMAVARISORPKAFEFVSKLYTDLTK 240
Db      205 KIDELDEBKASSAKORLKASLOKFGERAFAKMAVARISORPKAFEFVSKLYTDLTK 264
QY      241 VHTCCGHDLLECADRADLAKYICENODSISSKLECCCKPLLEKSHCIAEYENDMPA 300
Db      265 VHTCCGHDLLECADRADLAKYICENODSISSKLECCCKPLLEKSHCIAEYENDMPA 324
QY      301 DLPSTLAADFEESKDKVKNYAEAKDVFLGMFLYETARRHDPYSVLLRLAKTYETLLEKC 360
Db      325 DLPSTLAADFEESKDKVKNYAEAKDVFLGMFLYETARRHDPYSVLLRLAKTYETLLEKC 384
QY      361 CAADPHCEYAKYFDEKPLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKPVQVST 420
Db      385 CAADPHCEYAKYFDEKPLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKPVQVST 444
QY      421 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCCTES 480
Db      445 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCCTES 504
QY      481 LVNRRCFSALFEDVETVYKFEFNAETFFHADICTLSEKROJKTOTALVELVKKHPRKAT 540
Db      505 LVNRRCFSALFEDVETVYKFEFNAETFFHADICTLSEKROJKTOTALVELVKKHPRKAT 564
QY      541 KEOLKAVMDFAAFVCKCKADDKETCFABEGRKLVAAASQAALGL 585
Db      565 KEOLKAVMDFAAFVCKCKADDKETCFABEGRKLVAAASQAALGL 609

```

RESULT 14 US-10-237-624-2

; Sequence 2, Application US/10237624
; Publication No. US20030082747A1

GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard

Fourtner, Alain
Guitton, Jean-Dominique
Jung, Gerard

Yeh, Patricia

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.

STREET: 500 Atcoia Road, 3C43

CITY: Collegeville

STATE: PA

COUNTRY: USA

ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: Macintosh

OPERATING SYSTEM: System 7.1

SOFTWARE: Word 5.1 (Patentln)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/237, 624

FILING DATE: 10-Sep-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/797, 689

FILING DATE: 31-JAN-1997

APPLICATION NUMBER: US 08/256, 927

FILING DATE: 28-JUL-1994

APPLICATION NUMBER: FR 92/01064

FILING DATE: 31-JAN-1992

APPLICATION NUMBER: PCT/FR93/00085

```

; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-624-2

```

Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.3e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 DAHSEVARIKFDLGEENKALVLIAPAOYLQCCPEEDHVKLVNEYTERAKTCVADSEAE 60
Db      25 DAHSEVARIKFDLGEENKALVLIAPAOYLQCCPEEDHVKLVNEYTERAKTCVADSEAE 84
QY      61 NCDKSLHTLFSGDKLCTVATLRETYGEMADCCAKOBERNECEFLQHKDNDPNI.PRLVREPV 120
Db      85 NCDKSLHTLFSGDKLCTVATLRETYGEMADCCAKOBERNECEFLQHKDNDPNI.PRLVREPV 144
QY      121 DWGCTAFHNEETFLKKYLYETARRRPYAPPELLFFAKRYKAAPFECOAADKAACLP 180
Db      145 DWGCTAFHNEETFLKKYLYETARRRPYAPPELLFFAKRYKAAPFECOAADKAACLP 204
QY      181 KIDELDEBKASSAKORLKASLOKFGERAFAKMAVARISORPKAFEFVSKLYTDLTK 240
Db      205 KIDELDEBKASSAKORLKASLOKFGERAFAKMAVARISORPKAFEFVSKLYTDLTK 264
QY      241 VHTCCGHDLLECADRADLAKYICENODSISSKLECCCKPLLEKSHCIAEYENDMPA 300
Db      265 VHTCCGHDLLECADRADLAKYICENODSISSKLECCCKPLLEKSHCIAEYENDMPA 324
QY      301 DLPSTLAADFEESKDKVKNYAEAKDVFLGMFLYETARRHDPYSVLLRLAKTYETLLEKC 360
Db      325 DLPSTLAADFEESKDKVKNYAEAKDVFLGMFLYETARRHDPYSVLLRLAKTYETLLEKC 384
QY      361 CAADPHCEYAKYFDEKPLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKPVQVST 420
Db      385 CAADPHCEYAKYFDEKPLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKPVQVST 444
QY      421 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCCTES 480
Db      445 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYTKCCTES 504
QY      481 LVNRRCFSALFEDVETVYKFEFNAETFFHADICTLSEKROJKTOTALVELVKKHPRKAT 540
Db      505 LVNRRCFSALFEDVETVYKFEFNAETFFHADICTLSEKROJKTOTALVELVKKHPRKAT 564
QY      541 KEOLKAVMDFAAFVCKCKADDKETCFABEGRKLVAAASQAALGL 585
Db      565 KEOLKAVMDFAAFVCKCKADDKETCFABEGRKLVAAASQAALGL 609

```

RESULT 15

US-10-153-064-133

; Sequence 133, Application US/10153064

; Publication No. US20020142814A1

GENERAL INFORMATION:

APPLICANT: Bell et al.

TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

FILE REFERENCE: PF556

CURRENT APPLICATION NUMBER: US/10/153, 064

CURRENT FILING DATE: 2002-05-24

PRIOR APPLICATION NUMBER: 60/293,212

;; PRIOR FILING DATE: 2001-05-25
;; NUMBER OF SEQ ID NOS: 137
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 133
;; LENGTH: 651
;; TYPE: PR
;; ORGANISM: Homo sapiens
US-10-153-064-133

Query Match 100.0%; Score 3103; DB 15; Length 651;
Best Local Similarity 100.0%; Pred. No. 9,1e-271;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDGEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 67 DAHKEVAHRRFKDGEENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 126
QY 61 NCDSLHTLFGDKICTATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVPREV 120
DB 127 NCDSLHTLFGDKICTATLRETYGEMADCCAKOEPERNECFLOHKDNPMLPRLVPREV 186
QY 121 DVMTAFHDNEETFLKKYLYEIARRHRYFYAPPELLFFAKRYKAFTGCCQAADKAACILP 180
DB 187 DVMTAFHDNEETFLKKYLYEIARRHRYFYAPPELLFFAKRYKAFTGCCQAADKAACILP 246
QY 181 KLDELDEGKAASSAKORLKCASLOKFGERAFAKANAVALSQRPPKAEPAYSKLVTDLT 240
DB 247 KLDELDEGKAASSAKORLKCASLOKFGERAFAKANAVALSQRPPKAEPAYSKLVTDLT 306
QY 241 VHTCCGHDLLCEADDRADLAKYICENODSISSKLKECCCKPILLEKSHCIAEVENDEMPA 300
DB 307 VHTCCGHDLLCEADDRADLAKYICENODSISSKLKECCCKPILLEKSHCIAEVENDEMPA 366
QY 301 DLPSLADFEVSKDVCNKYAEAKDVFGLMFLYETARRHPDYSVLLRLAKTYETTLK 360
DB 367 DLPSLADFEVSKDVCNKYAEAKDVFGLMFLYETARRHPDYSVLLRLAKTYETTLK 426
QY 361 CAADPHECYAKVDEEKPIVEEPONL IKONCELEEQLEGEYKFONALLVRYTKKVPQVST 420
DB 427 CAADPHECYAKVDEEKPIVEEPONL IKONCELEEQLEGEYKFONALLVRYTKKVPQVST 486
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHETKTPVSDRYTKCCTES 480
DB 487 PTLVEVSRLNGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHETKTPVSDRYTKCCTES 546
QY 481 LVNRRPCFSALEVDETYVVPKEFNATFTFHADICTLSSEKEROIKKOTALVELVKKHPRAT 540
DB 547 LVNRRPCFSALEVDETYVVPKEFNATFTFHADICTLSSEKEROIKKOTALVELVKKHPRAT 606
QY 541 KEQIKAVMDDFAAFEVERCKKADKKECTCFAEGRKLVAAQAALGI 585
DB 607 KEQIKAVMDDFAAFEVERCKKADKKECTCFAEGRKLVAAQAALGI 651

Search completed: July 22, 2003, 11:53:57
Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 22, 2003, 11:42:49 ; Search time 22 Seconds
(Without alignments) 2556.302 Million cell updates/sec

Title: US-09-833-118-18
Perfect score: 3103

Sequence: 1 DAHSEVAHRFKDLEGENFK.....TCFAEGRKLVASQAALGL 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_73:*
2: pir1:*
3: pir2:*
4: pir3:*
5: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	609	1	ABHUS serum albumin prec
2	2942	94.8	609	2	A47391 serum albumin prec
3	2620	84.4	608	2	S57632 serum albumin prec
4	2475.5	79.8	607	1	ABHOS serum albumin prec
5	2446.5	78.8	607	1	ABHOS serum albumin prec
6	2432.5	78.4	607	1	ABSHS serum albumin prec
7	2426	78.2	608	1	ABRTS serum albumin prec
8	2411.5	77.7	605	1	ABPGS serum albumin prec
9	2387	76.9	609	2	JC5838 albumin - Mongolia
10	1861	60.0	453	2	A05139 serum albumin - mo
11	1557.5	50.2	615	1	ABCHS serum albumin prec
12	1253.5	40.4	609	2	JC4258 alpha-fetoprotein
13	1249.5	40.3	609	1	FPPO alpha-fetoprotein
14	1242.5	40.0	609	1	FPPO alpha-fetoprotein
15	1205	38.8	607	1	ABXL72 74k albumin prec
16	1181.5	38.1	265	2	I46968 albumin - dog (fra
17	1175.5	37.9	608	1	ABXL68 68k serum albumin
18	1084	34.9	605	1	FPMS alpha-fetoprotein
19	1064	34.4	611	1	FPPT alpha-fetoprotein
20	1055	34.0	599	1	A54906 albumin precursor -
21	928.5	29.9	614	2	S59517 serum albumin prec
22	928	29.9	608	2	A53195 albumin precursor -
23	747.5	23.1	608	1	ABONS1 serum albumin 1 pr
24	742.5	23.9	608	1	ABONS2 serum albumin 2 pr
25	699	22.5	382	2	A37253 serum albumin - bu
26	440.5	14.2	1423	1	S27941 serum albumin - se
27	386	12.4	474	1	VYHVD vitamin D-binding
28	385	12.4	476	1	VYRVD vitamin D-binding
29	372	12.0	472	1	A35327 vitamin D-binding

30	184	5.9	1819	2	A71928	cag island protein
31	184	5.9	1927	2	G64585	cag pathogenicity
32	141.5	4.6	1560	2	T30282	calcium-binding pr
33	137.5	4.4	1348	2	AG2558	hypothetical prote
34	134.5	4.3	1004	2	JC2221	major surface glyco
35	134	4.3	1083	2	JC2300	cell surface glyco
36	134	4.3	1780	2	T11722	hypothetical prote
37	133.5	4.3	1390	2	S51364	sperm tail-specific
38	132.5	4.3	1070	2	T06733	kinesin homolog F2
39	131	4.2	1076	2	JC2217	major surface glyco
40	130	4.2	1175	2	D35815	myosin heavy chain
41	129	4.2	1017	2	PC4035	cell-cycle-depende
42	128.5	4.1	1051	2	JC4091	glycoprotein A - p
43	128	4.1	1005	2	A64465	hypothetical prote
44	128	4.1	3225	2	I52300	giantin - human
45	126.5	4.1	1927	2	A59236	embryonic muscle m

ALIGNMENTS

RESULT 1

ABHOS serum albumin precursor [validated] - human

N:Alternate names: preproalbumin
N:Contains: Kinetensin

C:Species: Homo sapiens (man)

C>Date: 29-Jul-1981 #sequence, revision 31-Jan-1997 #text, change 17-Mar-2000

C:Accession: A93743; A93936; I59427; I59286; I59313; G01747; S55314; A91420; S06422;

R:Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najjarian, R.C.; See

Nucleic Acids Res. 9, 6103-6114, 1981

A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia c

A:Reference number: A93743; MUID:82081882; PMID:6171778

A:Accession: A93743

A:Molecule type: mRNA

A:Residues: 1-419, 'K', 421-609 <LAW>

A:Cross-references: EMBL:V00495; GB:J00078; GB:U00132; GB:L00133; NID:928591; PIDN:CA

R:Dugalczyk, A.; Law, S.W.; Denison, O.E.

A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A:Reference number: A93936; MUID:82105994; PMID:6275391

A:Accession: A93936

A:Molecule type: mRNA

A:Residues: 1-120, 'G', 122-609 <DUG>

A:Cross-references: EMBL:V00494; NID:928589; PIDN:CA23753.1; PID:928590

R:Uranio, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.

J. Biol. Chem. 261, 3244-3251, 1986

A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions a

A:Reference number: I39427; MUID:86140099; PMID:2419329

A:Accession: I39427

A:Molecule type: not shown

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-26 <URA>

A:Cross-references: GB:M13075; NID:9178330; PIDN:AA51688.1; PID:9553173

R:Watkins, S.; Madison, J.; Galliano, M.; Minichotti, L.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994

A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian faml

A:Reference number: I59286; MUID:94181575; PMID:8134387

A:Accession: I59286

A:Molecule type: DNA

A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Cross-references: GB:S69192; NID:9546032; PIDN:AA30282.1; PID:9546033

A>Note: This frame-shift variant, designated albumin Roma, leads to analbuminemia

R:Madison, J.; Galliano, M.; Watkins, S.; Minichotti, L.; Porta, F.; Rossi, A.; Putna

Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994

A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carbox

A:Reference number: I59313; MUID:94294404; PMID:8022807

A:Accession: I59313

A:Molecule type: DNA

A>Status: translated from GB/EMBL/DDBJ

A:Residues: 589-590, 'ALPRRVNLLQVRLP' <MAD>

A:Cross-references: GB:S70799; NID:9547231; PIDN:AA31177.1; PID:9547232

A:Note: this frame-shift variant is designated albumin Bazzano: four additional variants
 R:Menaya, J.; Parrilla, R.; Ayuso, M.S.
 Submitted to the EMBL Data Library, March 1995
 A:Reference number: G08292
 A:Accession: G01747
 A:Status: translated from GR/EMBL/DBD
 A:Molecule type: mRNA
 A:Residues: 1-120,'G',122-455 <MEN>
 A:Cross-references: EMBL:U25961; NID:q763428; PIDN:AAA64922.1; PID:q763431
 R:Riedge, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
 Biochem. J. 308, 321-325, 1995
 A:Title: Enzymotelytic processing of recombinant proalbumin variants by the yeast *Kex2*
 A:Reference number: S5314; MUID:95275251; PMID:7755581
 A:Accession: S5314
 A:Molecule type: protein
 A:Residues: 19-27 <LED>
 R:Mejoun, B.; Moravsek, L.; Kostka, V.
 FEBS Lett. 58, 134-137, 1975
 A:Title: Complete amino acid sequence of human serum albumin.
 A:Reference number: A91420; MUID:76187907; PMID:1225573
 A:Accession: A91420
 A:Molecule type: protein
 A:Residues: 25-117,'E',120-154,'Q',156-193,'E',195-387,'H',389-390,'Y',392-393,'A',395-
 R:Rocher, U.; Spittler, G.; Tripler, D.
 Justus Liebig's Ann. Chem. 9, 881-884, 1988
 A:Title: Isolation and structure elucidation of middle-molecular weight peptides from ur
 A:Reference number: S06422
 A:Note: this paper is in German, with an English abstract
 A:Accession: S06422
 A:Molecule type: protein
 A:Residues: 25-48 <ROE>
 R:Fluch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
 Arch. Biochem. Biophys. 305, 595-599, 1993
 A:Title: Mass spectrometric identification of modifications to human serum albumin treat
 A:Reference number: S36882; MUID:93384321; PMID:8373198
 A:Accession: S36882
 A:Molecule type: protein
 A:Residues: 45-67;141-160;311-337;469-490;570-581 <FIN>
 R:Kausler, E.; Spittler, G.
 Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
 A:Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mitteilmo
 A:Reference number: S17599; MUID:92126241; PMID:1772598
 A:Accession: S17599
 A:Molecule type: protein
 A:Residues: 25-54;354-357;431-447 <KAU>
 A:Note: 49-Leu was also found
 R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
 J. Immunol. 143, 1680-1684, 1989
 A:Title: Structures of histamine-releasing peptides formed by the action of acid proteas
 A:Reference number: A45800; MUID:89341406; PMID:2474609
 A:Accession: A45800
 A:Molecule type: protein
 A:Residues: 166-173 <CAR>
 R:Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa
 Biochem. Biophys. Res. Commun. 136, 983-988, 1986
 A:Title: The amino acid sequence of kineinsin, a novel peptide isolated from pepsin-tre
 A:Reference number: A03239; MUID:86242180; PMID:3087352
 A:Accession: A03239
 A:Molecule type: protein
 A:Residues: 166-173,'U' <MOC>
 R:Galliano, M.; Minichotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S
 Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
 A:Title: Mutations in genetic variants of human serum albumin found in Italy.
 A:Reference number: A38255; MUID:91062352; PMID:2247440
 A:Accession: C38255
 A:Molecule type: protein
 A:Residues: 76-111 <GAL1>
 A:Accession: B38255
 A:Molecule type: protein
 A:Residues: 82-105,'K',107-110 <GAL2>
 A:Note: this variant is designated albumin Vibo Valentia
 A:Accession: A38255
 A:Molecule type: protein

A:Residues: 76-83,'K',85-106 <GAL3>
 A:Note: this variant is designated albumin Torino
 R:Minichotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
 Eur. J. Biochem. 214, 437-444, 1993
 A:Title: The structural characterization and bilirubin-binding properties of albumin
 A:Reference number: S33298; MUID:95292504; PMID:8513793
 A:Accession: S33298
 A:Molecule type: protein
 A:Residues: 255-263,'E',265-281 <MIN1>
 A:Note: this variant is designated albumin Herborn
 R:Minichotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Por
 Biochim. Biophys. Acta 1119, 23-28, 1992
 A:Title: Two allolbumins with identical electrophoretic mobility are produced by dif
 A:Reference number: S21078; MUID:92190239; PMID:1347703
 A:Accession: S21078
 A:Molecule type: protein
 A:Residues: 354-356,'K',358-378 <MIN2>
 A:Note: this variant is designated albumin Sondrio; another variant Paris-2 is report
 R:He, X.M.; Carter, D.C.
 Nature 358, 209-215, 1992
 A:Title: Atomic structure and chemistry of human serum albumin.
 A:Reference number: A46756; MUID:92334427; PMID:1630489
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms
 R:Brown, J.R.; Shockley, P.; Behrens, P.Q.
 In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp. 23-
 A:Reference number: A94442
 A:Contents: annotation; three-dimensional structure and disulfide bonds
 R:Saber, M.A.; Stockbauer, P.; Moravsek, L.; Mejoun, B.
 Collect. Czech. Chem. Commun. 42, 564-579, 1977
 A:Title: Disulfide bonds in human serum albumin.
 A:Reference number: A90930
 A:Contents: annotation; disulfide bonds
 R:Jacobsen, C.
 Biochem. J. 171, 453-459, 1978
 A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity bindi
 A:Reference number: A90289; MUID:78186630; PMID:656055
 A:Contents: annotation; bilirubin-binding site
 R:Peterson, T.; Reed, R.G.
 In Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjoblom, I., eds., 11-
 A:Title: Serum albumin: conformation and active sites.
 A:Reference number: A94408
 A:Contents: annotation; binding sites
 R:Harper, M.E.; Dugalczyk, A.
 Am. J. Hum. Genet. 35, 565-572, 1983
 A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein ge
 A:Reference number: A90028; MUID:83279982; PMID:6192711
 A:Contents: annotation; gene position
 R:Walker, J.E.
 FEBS Lett. 66, 173-175, 1976
 A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic ac
 A:Reference number: A46755; MUID:76257808; PMID:955075
 A:Contents: annotation
 A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic a
 R:Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
 FEBS Lett. 298, 266-268, 1992
 A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phos
 A:Reference number: A56294; MUID:92183881; PMID:1544460
 A:Contents: annotation
 A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; 1
 atase activity
 C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
 in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (we
 C:Comment: A large number of variants of human serum albumin have been described.
 C:Genetics:
 A:Gene: GDB:ALB
 A:Cross-references: GDB:118990; OMTM:103600
 A:Map position: 4q11-q13
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyri
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status experimental <PRO>
 F:25-609/Product: serum albumin #status experimental <PT>
 F:29-202/Domain: serum albumin repeat homology <SA1>

F:166-174/Product: kinetensin #status experimental <KIP>
 F:221-394/Domain: serum albumin repeat homology <SA2>
 F:413-592/Domain: serum albumin repeat homology <SA3>
 F:27/Binding site: copper (His) #status predicted
 F:77-86,99-115,114-125,148-193,192-201,224-270,265-277,289-303,302-313,340-385,384-393,4
 F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 3103; DB 1; Length 609;
 Best Local Similarity 100.0%; Pred. No. 7,4e-198;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHFRFDLGEENFKALVLIAPAYILOQCPEDHVKLVNTEFEAKTCVADESA 60
 DB 25 DAHKEVAHFRFDLGEENFKALVLIAPAYILOQCPEDHVKLVNTEFEAKTCVADESA 84
 QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEERNECELOHKDNDPMLPLVREPV 120
 DB 85 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEERNECELOHKDNDPMLPLVREPV 144
 QY 121 DVMCTAFHNEETFLKKYLETARRHPYPAPELLFFAKRYKAFTCCOADAACCLIP 180
 DB 145 DVMCTAFHNEETFLKKYLETARRHPYPAPELLFFAKRYKAFTCCOADAACCLIP 204
 QY 181 KLDELDEBKASSAKORLKCASLOKGERAFKAMAVARLSORPKAFPAEVSRLVDTLK 240
 DB 205 KLDELDEBKASSAKORLKCASLOKGERAFKAMAVARLSORPKAFPAEVSRLVDTLK 264
 QY 241 VTECHGDLLECCADRADLAKYICENODSISSKLECCCKPLEKSHCIAEYENDMPA 300
 DB 265 VTECHGDLLECCADRADLAKYICENODSISSKLECCCKPLEKSHCIAEYENDMPA 324
 QY 301 DPLSLAADVYESKDVCKNAEAKDVPLGMFLYEYARRHPDYSVLLRLAKYETTLK 360
 DB 325 DPLSLAADVYESKDVCKNAEAKDVPLGMFLYEYARRHPDYSVLLRLAKYETTLK 384
 QY 361 CAADPHCEYAKVFDEKFLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKPVYST 420
 DB 385 CAADPHCEYAKVFDEKFLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKPVYST 444
 QY 421 PTLVEYSRLGKVGSKCKHPKAPKPCADLYLVLNOLCVLHEKTPVSDRYTKCTES 480
 DB 445 PTLVEYSRLGKVGSKCKHPKAPKPCADLYLVLNOLCVLHEKTPVSDRYTKCTES 504
 QY 481 LVNRRPCFSALVEDETYVKEFNAETFTHADICTISEKERQIKKOTALVELYKHKPKAT 540
 DB 505 LVNRRPCFSALVEDETYVKEFNAETFTHADICTISEKERQIKKOTALVELYKHKPKAT 564
 QY 541 KEOLKAVMDFAAFVEKCKKADDEKTCFAEBEGKTLVAASQALGL 585
 DB 565 KEOLKAVMDFAAFVEKCKKADDEKTCFAEBEGKTLVAASQALGL 609

RESULT 2

A47391
 serum albumin precursor - rhesus macaque
 C:Species: Macaca mulatta (rhesus macaque)
 C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
 C:Accession: A47391
 R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwalet, J.; Putnam, F.
 Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
 A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in b1/b2
 A:Reference number: A47391; MUID:93211971; PMID:8460152
 A:Contents: B/B homozygote
 A:Accession: A47391
 A:Status: preliminary
 A:Molecule type: mRNA; protein
 A:Residues: 1-600 <MAT>
 A:Cross-references: GB:M90463; NID:9342294; PIDN:AAA36906.1; PID:9342295
 A:Experimental source: liver
 A:Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIPI:128281)
 C:Superfamily: serum albumin; serum albumin repeat homology
 F:21-194/Domain: serum albumin repeat homology <SA1>
 F:213-386/Domain: serum albumin repeat homology <SA2>

F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 94.8%; Score 2962; DB 2; Length 600;
 Best Local Similarity 93.5%; Pred. No. 3.3e-187;
 Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHKEVAHFRFDLGEENFKALVLIAPAYILOQCPEDHVKLVNTEFEAKTCVADESA 60
 DB 17 DAHKEVAHFRFDLGEENFKALVLIAPAYILOQCPEDHVKLVNTEFEAKTCVADESA 76
 QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEERNECELOHKDNDPMLPLVREPV 120
 DB 77 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEERNECELOHKDNDPMLPLVREPV 136
 QY 121 DVMCTAFHNEETFLKKYLETARRHPYPAPELLFFAKRYKAFTCCOADAACCLIP 180
 DB 137 DVMCTAFHNEETFLKKYLETARRHPYPAPELLFFAKRYKAFTCCOADAACCLIP 196
 QY 181 KLDELDEBKASSAKORLKCASLOKGERAFKAMAVARLSORPKAFPAEVSRLVDTLK 240
 DB 197 KLDELDEBKASSAKORLKCASLOKGERAFKAMAVARLSORPKAFPAEVSRLVDTLK 256
 QY 241 VTECHGDLLECCADRADLAKYICENODSISSKLECCCKPLEKSHCIAEYENDMPA 300
 DB 257 VTECHGDLLECCADRADLAKYICENODSISSKLECCCKPLEKSHCIAEYENDMPA 316
 QY 301 DPLSLAADVYESKDVCKNAEAKDVPLGMFLYEYARRHPDYSVLLRLAKYETTLK 360
 DB 317 DPLSLAADVYESKDVCKNAEAKDVPLGMFLYEYARRHPDYSVLLRLAKYETTLK 376
 QY 361 CAADPHCEYAKVFDEKFLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKPVYST 420
 DB 377 CAADPHCEYAKVFDEKFLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKPVYST 436
 QY 421 PTLVEYSRLGKVGSKCKHPKAPKPCADLYLVLNOLCVLHEKTPVSDRYTKCTES 480
 DB 437 PTLVEYSRLGKVGSKCKHPKAPKPCADLYLVLNOLCVLHEKTPVSDRYTKCTES 496
 QY 481 LVNRRPCFSALVEDETYVKEFNAETFTHADICTISEKERQIKKOTALVELYKHKPKAT 540
 DB 497 LVNRRPCFSALVEDETYVKEFNAETFTHADICTISEKERQIKKOTALVELYKHKPKAT 556
 QY 541 KEOLKAVMDFAAFVEKCKKADDEKTCFAEBEGKTLVAASQALGL 583
 DB 557 KEOLKAVMDFAAFVEKCKKADDEKTCFAEBEGKTLVAASQALGL 599

RESULT 3

S57632
 serum albumin precursor - cat
 C:Species: Felis silvestris catus (domestic cat)
 C>Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
 C:Accession: J04660; S57632
 R:Hilger, C.; Grigioni, F.; Hentges, F.
 Gene 169, 295-296, 1996
 A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
 A:Reference number: J04660; MUID:96194824; PMID:8647469
 A:Accession: J04660
 A:Molecule type: mRNA
 A:Residues: 1-608 <H12>
 A:Cross-references: EMBL:X84842; NID:9886484; PIDN:CAA59279.1; PID:9886485
 A:Experimental source: liver
 C:Comment: This protein is the major protein component in plasma. It functions as a m
 eln has 35 conserved cysteine residues.
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: liver; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status predicted <PRP>
 F:25-608/Product: serum albumin #status predicted <MAT>
 F:29-202/Domain: serum albumin repeat homology <SA1>
 F:221-394/Domain: serum albumin repeat homology <SA2>
 F:413-592/Domain: serum albumin repeat homology <SA3>

A:Residues: 1-32 <MAG>
 R:Haieh, J.C.; Lin, F.P.; Tam, M.F.
 Anal. Biochem. 170, 1-8, 1988
 A:Title: Electrophoretic onto glass-fiber filter from an analytical isoelectrofocusing g
 A:Reference number: A60808; MUID:88267456; PMID:3389500
 A:Accession: B60808
 A:Molecule type: protein
 A:Residues: 25-41 <HSI>
 R:Strawich, E.; Glincher, M.J.
 Eur. J. Biochem. 191, 47-56, 1990
 A:Title: Tooth enamelins identified mainly as serum proteins. Major 'enamelin' is albu
 A:Reference number: S10780; MUID:90336641; PMID:2379503
 A:Accession: S10780
 A:Molecule type: protein
 A:Residues: 25-41, 'H', 43-57, 59-64 <SNR>
 R:Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mittra, S.P.
 J. Immunol. 143, 1680-1684, 1989
 A:Title: Structures of histamine-releasing peptides formed by the action of acid proteas
 A:Reference number: A45800; MUID:89341406; PMID:2474609
 A:Accession: D45800
 A:Molecule type: protein
 A:Residues: 163-172 <CAR>
 R:Carraway, R.E.; Mittra, S.P.; Cochrane, D.E.
 J. Biol. Chem. 262, 5968-5973, 1987
 A:Title: Structure of a biologically active neurotensin-related peptide obtained from pe
 A:Reference number: A26693; MUID:87194805; PMID:2437111
 A:Accession: A26693
 A:Molecule type: protein
 A:Residues: 165-172, 'L' <CA2>
 R:Reed, R.G.; Putnam, F.W.; Peters Jr., T.
 Biochem. J. 191, 867-868, 1980
 A:Title: Sequence of residues 400-403 of bovine serum albumin.
 A:Reference number: A90309; MUID:82023364; PMID:7283978
 A:Accession: A90309
 A:Molecule type: protein
 A:Residues: 402-433 <REB>
 R:Brown, J.R.
 Fed. Proc. 34, 591, 1975
 A:Title: Structure of bovine serum albumin.
 A:Reference number: A91458
 A:Accession: A91458
 A:Molecule type: protein
 A:Residues: 25-41, 'H', 43-117, 'EO', 120-179, 181-189, 'E', 191-194, 'A', 196-213, 'T', 215-288, 'H
 R:Brown, J.R.
 submitted to the Atlas, April 1975
 A:Reference number: A94551
 A:Accession: A94551
 A:Molecule type: protein
 A:Residues: 190-195 <BR2>
 R:Brown, J.R.
 Fed. Proc. 33, 1389, 1974
 A:Reference number: A91457
 R:Werlen, R.C.; Olford, R.E.; Rose, K.
 Biochem. J. 302, 907-911, 1994
 A:Title: Preparation and characterization of novel substrates of insulin proteinase (EC
 A:Reference number: S55232; MUID:95031935; PMID:7945219
 A:Accession: S55232
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 529-536;569-572 <MER>
 C:Superfamily: serum albumin; copper binding; duplication; plasma
 C:Keywords: carrier protein; copper binding; duplication; plasma
 F:1-18/Domain: signal sequence #status experimental <PRO>
 F:19-24/Domain: propeptide #status experimental <PRO>
 F:25-607/Product: serum albumin #status experimental <MPY>
 F:29-201/Domain: serum albumin repeat homology <SA1>
 F:220-393/Domain: serum albumin repeat homology <SA2>
 F:412-591/Domain: serum albumin repeat homology <SA3>
 F:27/Binding site: copper (His) #status predicted
 F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392, 4

Query Match 78.8%; Score 2446.5; DB 1; Length 607;

Best Local Similarity 75.6%; Pred. No. 2, 1e-154;
 Matches 441; Conservative 71; Mismatches 70; Indels 1; Gaps 1;

QY 1 DARKSEVAFKFKLGEENKALVLAFAQYLQCCPEEDHVKLVNTEFAKTVADSEAE 60
 Db 25 DFKRSEAFKFKLGEENKALVLAFAQYLQCCPEEDHVKLVNTEFAKTVADSEAE 84
 QY 61 NCDKSLHTLEGGDKLVATRETYGMAADCAQOEERNCFLQHDNDNPLRLRPEV 120
 Db 85 GCKSLHTLEGGDKLVATRETYGMAADCAQOEERNCFLQHDNDNPLRLRPEV 143
 QY 121 DVACTAFHDEETFLKLYEIAIRRRPYFAPLLEFAKKAFTCCQAADKAALP 180
 Db 144 NTLCDERKADKFKWKKLYEIAIRRRPYFAPLLEFAKKAFTCCQAADKAALP 203
 QY 181 KLELADDEKASSAKQRLKASQKGERAFKAWARLSQRPKAEFAVSKLVYDLTK 240
 Db 204 KITMREKVLASSARQRLKASQKGERAFKAWARLSQRPKAEFAVSKLVYDLTK 263
 QY 241 VHECCGHDLECCADRADLAKYICENQDSSSKLECKPPLKESHCIAEYNDMPA 300
 Db 264 VHECCGHDLECCADRADLAKYICENQDSSSKLECKPPLKESHCIAEYNDMPA 323
 QY 301 DLPSLADEYFESKDVCKNTAEAKDVEFGMELYEYARRHPDYVLLRLAKYETLEKC 360
 Db 324 NLPLADAFEDKDVCKNTAEAKDVEFGMELYEYARRHPDYVLLRLAKYETLEKC 383
 QY 361 CAADPHECYAKYFDEFKPLVEEPQNLKONCELFGQGEYKFNALVRYTKKPVQST 420
 Db 384 CAKDDEPHACYSTVEFKLHLVDEPQNLKONCELFGQGEYKFNALVRYTKKPVQST 443
 QY 421 PTLVEYSRNLGKYSCKCKRPEAKRMPCAEDYLSVLNQCIVHEPTVSDRYTKCTES 480
 Db 444 PTLVEYSRNLGKYSCKCKRPEAKRMPCAEDYLSVLNQCIVHEPTVSDRYTKCTES 503
 QY 481 LVNRPCFSALVEDEYTPKEFNAETFTFADICTSEKROKOTALELKHPRPKAT 540
 Db 504 LVNRPCFSALVEDEYTPKEFNAETFTFADICTSEKROKOTALELKHPRPKAT 563
 QY 541 KEOLKAVMDPAFVEKCKADKCEKFAEKGKLVAAQAL 583
 Db 564 EEOCLKTVMENFVAFVNDKCAADDEKACFAVEGPKLVSTOTAL 606

RESULT 6 ABSHS

serum albumin precursor - sheep
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: S06936
 R:Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
 Nucleic Acids Res. 17, 10495, 1989
 A:Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
 A:Reference number: S06936; MUID:90098888; PMID:2602160
 A:Accession: S06936
 A:Molecule type: mRNA
 A:Residues: 1-607 <BRO>
 A:Cross-references: EMBL:X17055; NID:g1386; PID:CAA3903.1; PID:g1387
 A:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds coppe
 teroid hormones (weak bonds with these hormones promote their transfer across the mem
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status predicted <PRO>
 F:25-607/Product: serum albumin #status predicted <MAP>
 F:29-201/Domain: serum albumin repeat homology <SA1>
 F:220-393/Domain: serum albumin repeat homology <SA2>
 F:412-591/Domain: serum albumin repeat homology <SA3>
 F:27/Binding site: copper (His) #status predicted
 F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-39

Query Match 78.4%; Score 2432.5; DB 1; Length 607;

Best Local Similarity 75.0% ; Pred. No 18e-153;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

OY 1 DAHKEVAHRFKDLGEENFKAFLVLAFAOYLQOCPEEDHWKLVNEYTEFAKTGVADESA 60
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 25 DTRKSEIAHFBNLDGEEFGVLIAFSYOLOCCPFDEHKVKLVELTEFAKTCADESNA 84

OY 61 NCDKSLLTLGLDKLCYTATLTRETYIGEMADCCAQOEPRNRCFIQHNDNNLRPLRVREV 120
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 85 GCCKSLHTLTLGGDELCKCATLTRETGYGDMAKDCEQEOPERNRCFNLNHKKDSPLDRL - KREP 143

OY 121 DVMTAFTHDNDETFPKLYLEIARHPYPAPAEELLFPFAKYTKAAFTGCCAAMAKCLIP 180
 | : | : | : | : | | | | | : | | | | | : | | | | | : | | | | |
Db 144 DTLSAEFRKADEKRTKGWYILEVARRRHYPTAPELLITANKYNGVFQCDAEKRGAQLP 203

OY 181 KLDELDEGRKASAKORLKCASLQKFGERAFKAMAVALRSOREPKAEFAEVSKLVTDLTR 240
 | : | : | : | : | | | | | : | | | | | : | | | | | : | | | | |
Db 204 KIDMAREKVLASSARQRRLCASIQKFGERALKMASVARLSQRPKPADFDTVTKIYDCLR 263

OY 241 VHECCCGDLLLECADDPAIDLAKTICENODSISKLKCCCKPILLESKHCIAEVNDEMRA 300
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 264 VHKECCHGDDLEECADDPADLAKTICTHQDALSSKLKCCCKPYLEKSHCIAEVDKAVPE 323

OY 301 DLPLEADFEVESKVCKNYEAARKDFGMFLFYEARRHDPYSVLLIRLKITETTELKC 360
 : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 324 NLPLPLTDFADBDKCVCKNTQDAKDYLGSPFLYEISRHHPEYAASVLRLLAKEATIELDC 363

OY 361 CAADPHCYAKVEDEFEPRLFVEEPONLIKONCELFEOLGEGKQNALLIVRYTKRPOVST 420
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 384 CAKEDPHACATATVPDKLHLVDPEPNLIKKNCELFKHGEYGPMNALIVRYTKRAPOVST 443

OY 421 PTVLEVSRNLKGVSKSCCKHPEAKRMPCAEYLSVINOLCVLHEKTPVSVDRTKCTES 480
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 444 PTVEISISLIKGVTGCKCACRESERMPTCEDYLSLINRLCVLHEKTPVSEKVTCTES 503

OY 481 LVNRPFCSALEVEDTVPKFNAAFETPPHADICTLSSEKROIKOTALVELVHKRKAT 540
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 504 LVNRRPFCSDTLLETITYPKRFDEKFTFHADICTLPDTEKOKOTALVELLKKRKAT 563

OY 541 KEQLKAWMDPFAEFVKCCAKADKETCFABEGKKLVIASAQAAL 583
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 564 DEQLKTWMENVAFVAVDKCCAADKDGCGVLEGPRPLASTQAAL 606

RESULT 7

A:BTS
seum albumin precursor - rat
N:Alternate names: preproalbumin
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-May-1979 #sequence revision 31-May-1979 #text change 22-Jun-1999
C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R:Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A>Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A:Reference number: A93872; MUID:81223722; PMID:7017712
A:Accession: A93872
A:Molecule type: mRNA
A:Residues: 1-608 <SNR>
A:Cross-references: GB:V01222; GB:T00698; NID:955627; PIDN:CAA24532.1; PID:955628
R:Straus, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A>Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis
A:Reference number: A92211; MUID:77249657; PMID:893447
A>Note: cleavages during protein maturation
A:Accession: A92211
A:Molecule type: protein
A:Residues: 1-38 <STR>
R:Isemura, S.; Ikemaka, T.
J. Biochem. 83, 35-48, 1978
A>Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
A:Reference number: A91946; MUID:78109429; PMID:564345
A:Accession: A91946
A:Molecule type: protein

A:Residues: 25-222 <IS1>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 79, 1183-1196, 1976
A:Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amin
A:Reference number: A91940; MUID:76260153; PMID:956149
A:Accession: A91940
A:Molecule type: protein
A:Residues: 223-288:572-608 <IS2>
A:Note: 262-Ileu was also found
R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A:Title: Copper(II)-binding ability of human alpha-fetoprotein.
A:Reference number: A90758; MUID:79001617; PMID:80265
A:Contents: annotation; copper binding
R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1969
A:Title: Structures of histamine-releasing peptides formed by the action of acid prot
A:Reference number: A45800; MUID:89341406; PMID:2474609
A:Accession: C45800
A:Status: preliminary
A:Molecule type: protein
A:Residues: 166-173 <CAR>
R:Heard, J.
Mol. Cell. Biol. 7, 2425-2434, 1987
A:Title: Determinants of rat albumin promoter tissue specificity analyzed by an impro
A:Reference number: 157621; MUID:87286876; PMID:3475566
A:Accession: 157621
A:Status: preliminary
A:Molecule type: DNA
A:Cross-references: GB:M16825; NID:q202828; PIDN:AAA40712.1; PID:g554412
A:Residues: 1-5 <RES>
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status experimental <SIG>
F:19-24/Domain: propeptide #status experimental <PRO>
F:25-608/Product: serum albumin #status experimental <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (his) #status experimental
F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-399

Query Match 78.2%; Score 2426; DB 1; Length 608;
Best Local Similarity 73.4%; Pred. No. 4.7e-153;
Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;

QY 1 DAHSEVNHRRKDDGEENFKALVLIARAOYLQDCPFEDHVKLVNVEYFAKTCYADESAE 60
DB 25 EAHNSEIHRKRKDDGEQHFVKLVLIANSQYLQCKRYEHIKLVGEVDFAKTCVADENAE 84
QY 61 NCDKSLHATLFEDDKLCTVATLTRETYGEMADCCAKOBERNECFLOHKDDNPMLPLVPREV 120
DB 85 NCDKSIHILTFEDDKLCAIPKLRDNYGELADCCAKOBERNECFLOHKDDNPMLPPFORPEA 144
QY 121 DWMCAPFHDNETFTFKKTVLEIARHPYFAVPELLEFFAKRYKKAFTTECCQADRAACILP 180
DB 145 EAMCSTFSENPSTSEFGHILVEVARRHPRFTYAPPELLIYIAEKINEVLITQCTTSDDAACILTP 204
QY 181 KLDELIRDGKASAKORLKASLQKFGERRAKANAVARLISORFPAKFAEYKSLVTDLT 240
DB 205 KLDNAKKEKALVAANARQKCKSMQRFGRRAKANAIVAMSGRFPNAPFAETITKLATDVTK 264
QY 241 VHTTECHGDDLLLECADDRAADLAKIYICENODSISSKLKECCERPLLEKSHICIAEVENDENPA 300
DB 265 INKECHCHDILLECCADRAELAKMYCENONATISSKLOACDCKPVLQSOCLAEETHDNLPA 324
QY 301 DLPSTAAAFVESKDKCKNYAEAKDVEFGMPLFYEARHRPDVSVYLLRLAKYETLEK 360
DB 325 DLPSTAAAFVEDEKCKNYAEAKDVEFGTFLFYESSRRHPDVSIVSILLRLAKYETLEK 384
QY 361 CAAPPHCYKAVFEEFKPLVEEPONLLKONCELEFQJGEKRFONALLVRYTKRVPOYST 420
DB 385 CAEGPPKCYGVGLAEPDPLVEEPRKLVKTNCELEKIGEGFGQNAVLVRYTQAPQYST 444

Db 330 LPSEVLEYEDKESVCSFEAGHDAFMAFEVYESSRRHPEFSLQIMRIAGYESLEKCC 389
 QY 362 AADPHCEYAKVDEFEKPLVEEPONLTKONCELFEQGEYKFPONALIVRTKVPQVSTP 421
 Db 390 KTRNPACVYANAOEQLNQHKETQDVYKTKCDLHGHGADFLKSLIRTKKMPQVPTD 449
 QY 422 TLVEVSRLNKVSGCKCKHPEAKRMPCAEDYLSVNLQCLVLEHKEFPVSDRYTKCCSESL 481
 Db 450 LLEETGCKMTTIGTKCCQQLQEDRRMACSEGLYSIVHDTCRKQETPTINDNVSCQCSQLY 509
 QY 482 VNRRPCEALEVEDETVYPRKEFNAETFEHADICTLSEKROIKQALVELVKKPRATK 541
 Db 510 ANRRPCTAGVDTYKTPPPFPNPFDMFSFDEKLSAPAEEREVGOMKLLINLRKPKQMT 569
 QY 542 EOLKAWDDFAAFVEKCKKADKDETCFAEGEKKLVAAASOALGL 585
 Db 570 EOLKTTADGTFAMVDCCKCOSDINTCGEGEAGNLIVOSRATLGI 613

RESULT 12

JCA4258
 alpha-fetoprotein precursor - chimpanzee
 C:Species: Pan troglodytes (chimpanzee)
 C:Date: 27-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 20-Aug-1999
 C:Accession: JCA4258
 R:Nishio, H.; Gibbs, P.E.M.; Minghetti, P.F.; Zielinski, R.; Dugalczyk, A.
 Gene 162, 213-220, 1995
 A:Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to h
 A:Reference number: JCA4258; MUID:96032345; PMID:7357431
 A:Accession: JCA4258
 A:Molecule type: DNA
 A:Residues: 1-609 <NDS>
 A:Cross-references: GB:021916; NID:9841311; PIDN:AA91641.1; PID:9841312
 C:Comment: This protein is a plasma protein produced in the fetal and neonatal liver and
 o similar properties and structure.
 C:Genetics:
 A:Gene: afp
 A:Map position: 3p
 A:Insertions: 29/1: 46/2: 90/3: 161/2: 205/3: 238/2: 281/3: 353/2: 397/3: 430/2: 476/3: 551
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: glycoprotein
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-609/Product: alpha-fetoprotein #status predicted <MAT>
 F:29-202/Domain: serum albumin repeat homology <SA1>
 F:221-394/Domain: serum albumin repeat homology <SA2>
 F:413-592/Domain: serum albumin repeat homology <SA3>
 F:42,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.4%; Score 1253.5; DB 2; Length 609;
 Best Local Similarity 40.1%; Pred. No. 1.9e-75;
 Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;

QY 3 HKE-----VAHRRKIDGEEFKALVILIAFQYIQQCPFEDHVKLVNEVEFAKTVAD 57
 Db 22 HRNEVGIASILDYDQCTAEINLTDLAIIFPAQFOEATYKEVSMVADALTAIEKPGDE 81
 QY 58 SANCCKSLHTLFGDKICVATLRBYTGEMADCCAKPRRNCCELOHKNDP-NCRLV 116
 Db 82 QSGAGCLENOLPALLEEICRKEKTELEKGGH-SDCCSOSSEGRHMCFLAHKRPDPASIPFQ 140
 QY 117 RPEVDWCTAFHNDNEETFLKLYLEIARRHPYFAPELLEFAKRYKAFFECQOAAKAA 176
 Db 141 VPEPVASCVEYEDRETFPMKFIYEIARRHPILAPILLMAARYDKIIFSCCKAENAVE 200
 QY 177 CLIPKLELDEKCKASAKORLCASTLQFGERAFKAAVAVARLSQRPKAEPAVSKLV 236
 Db 201 CFQTKATVYKELEKRESILNQHCAYWKNNGTRTFQAIYTKLSQKFTKYNFEIOQLVL 260
 QY 237 DLKRVHEGCHGLLECADRADLAKYICENODSISKLECKECPKLLKSHCAEVD 296
 Db 261 DVAHVHHCGRGVDLDLQGEKIMSTYISQODPLSKITCECKLTTLTEGQCIHAEND 320
 QY 297 EMPADLPISLADEVESKDVCKNTAEAKDVFLGMEFLYEARRHDPDYSVLLRLAKTYET 356

Db 321 EKEPGLSPNLRFLGDRDFNQFSGSEKNIFLASVHEYSRRHPOLAVSVILRAVAKQOEL 380
 QY 357 LKCCAAADHECYAKFDEFKPLVEEPOLITKONCELFEQGEYKFPONALIVRTKVP 416
 Db 381 LKCFQTEPNLEQCDKKEELQYIOESQALARKSCGLFQKGEYIQNMFVAAYTKAP 440
 QY 417 QVSTPELVESRLNKVSGCKCKHPEAKRMPCAEDYLSVNLQCLVLEHKEFPVSDRYTKC 476
 Db 441 QLTSSLEMAITRKMAATAATCCQQLSEDKLACGEAADITIGHLCIHETTPVNPVGGC 500
 QY 477 CTESLVNRRPCEFSALVEDETVYPRKEFNAETFEHADICTLSEKROIKQALVELVKKR 536
 Db 501 CTSSYANRRPCEFSLVYDEVYVPAFSDKFIHFHKLQCAQVAGALQFMKQEPILNIVKOK 560
 QY 537 PKATKEDLKAVMDDFAAFVEKCKKADKDETCFAEGEKKLVAAASOALGL 585
 Db 561 POITEQLEAVIADFSGLLEKCCGOGROEVECFEFGCKLISKTRALGV 609

RESULT 13

FPHU
 alpha-fetoprotein precursor [validated] - human
 N:Alternate names: AFP; alpha-1-fetoprotein; alpha-fetoglobulin
 C:Species: Homo sapiens (man)
 C:Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 08-Dec-2000
 C:Accession: A26624; S37655; A93961; A91497; A23699; A61480; A90624; A90757; A93042;
 R:Gibbs, P.E.M.; Zielinski, R.; Boyd, C.; Dugalczyk, A.
 Biochemistry 26, 1332-1343, 1987
 A:Title: Structure, polymorphism, and novel repeated DNA elements revealed by a compl
 A:Reference number: A26624; MUID:87185438; PMID:2436651
 A:Accession: A26624
 A:Molecule type: DNA
 A:Residues: 1-609 <GIB>
 A:Cross-references: GB:M16110; NID:g773678; PIDN:AAB58754.1; PID:g178236
 R:McVey, J.H.; Michaelides, K.; Hansen, L.P.; Ferguson-Smith, M.; Tilgman, S.; Kruml
 Hum. Mol. Genet. 2, 379-384, 1993
 A:Title: A G->A substitution in an HNF I binding site in the human alpha-fetoprotein
 A:Reference number: S37655; MUID:93278385; PMID:7664942
 A:Accession: S37655
 A:Molecule type: DNA
 A:Residues: 1-28 <MCV>
 A:Cross-references: EMBL:Z19532; NID:g28527; PIDN:CAA79592.1; PID:g28528
 A:Note: the authors translated the codon TAT for residue 26 as Thr
 R:Moriaga, T.; Sakai, M.; Wegmann, T.G.; Tamaoki, T.
 Proc. Natl. Acad. Sci. U.S.A. 80, 4604-4608, 1983
 A:Title: Primary structures of human alpha-fetoprotein and its mRNA.
 A:Reference number: A93961; MUID:83273664; PMID:6192439
 A:Accession: A93961
 A:Molecule type: mRNA
 A:Residues: 1-609 <MOR>
 A:Cross-references: GB:U00077; NID:g311348; PIDN:CAA24758.1; PID:g31351
 R:Beattie, W.G.; Dugalczyk, A.
 Gene 20, 415-422, 1982
 A:Title: Structure and evolution of human alpha-fetoprotein deduced from partial sequ
 A:Reference number: A91497; MUID:83188778; PMID:6187626
 A:Accession: A91497
 A:Molecule type: mRNA
 A:Residues: 429-556 <BEA>
 A:Cross-references: GB:U00076
 R:Pucci, P.; Siciliano, R.; Malorni, A.; Martino, G.; Tecce, M.F.; Ceccarini, C.; Terr
 Biochemistry 30, 5061-5066, 1991
 A:Title: Human alpha-fetoprotein primary structure: a mass spectrometric study.
 A:Reference number: A23699; MUID:91242409; PMID:1709810
 A:Accession: A23699
 A:Molecule type: protein
 A:Residues: 19-45;60-97;102-107;122-184;187-249;255-489;507-609 <PUC>
 R:Tecce, M.F.; Terrana, B.; Giuliani, M.M.; Ceccarini, C.
 J. Nucl. Med. Allied Sci. 34, 213-216, 1990
 A:Title: Characterization of in vitro expressed human alpha-fetoprotein as highly rep
 A:Reference number: A61480; MUID:91225826; PMID:1709209
 A:Accession: A61480
 A:Molecule type: protein

A.Residues: 19-45:63-97;102-107;122-184;187-249;255-489;507-609 <TEC>
R.Yachnin, S.; Hsu, R.; Heinrikson, R.L.; Miller, J.B.
Biochem. Biophys. Acta 493, 418-428, 1977
A.Title: Studies on human alpha-fetoprotein. Isolation and characterization of monomeric
A.Reference number: A90624; MUID:77242506; PMID:70228
A.Accession: A90624
A.Molecule type: protein
A.Residues: 'S',20-22,'S',24-35 <YAC>
A.Note: dimeric and trimeric forms have been found in addition to the monomeric form
R.Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 37, 3663-3667, 1977
A.Title: Comparative chemical structure of human alpha-fetoproteins from fetal serum and
A.Reference number: A90757; MUID:78001670; PMID:71198
A.Accession: A90757
A.Molecule type: protein
A.Residues: 'S',20-30,'A',32-37,'A' <AOY>
R.Ruoslahti, E.; Plhko, H.; Vaheri, A.; Seppala, M.; Virolainen, M.; Kontinen, A.
Johns Hopkins Med. J. Suppl. 3, 249-255, 1974
A.Title: 20. Alpha fetoprotein: structure and expression in man and inbred mouse strains
A.Reference number: A93042; MUID:75018719; PMID:4138095
A.Accession: A93042
A.Molecule type: protein
R.Sakai, M.; Morinaga, T.; Urano, Y.; Watanabe, K.; Wegmann, T.G.; Tamacki, T.
J. Biol. Chem. 260, 5055-5060, 1985
A.Title: The human alpha-fetoprotein gene. Sequence organization and the 5' flanking reg
A.Reference number: A92520; MUID:85182629; PMID:2580830
A.Contents: annotation; gene, exons and introns
R.Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A.Title: Copper(II)-binding ability of human alpha-fetoprotein.
A.Reference number: A90758; MUID:79001677; PMID:80265
A.Contents: annotation; metal binding
R.Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 39, 3571-3574, 1979
A.Title: alpha-fetoprotein as a carrier protein in plasma and its bilirubin-binding abil
A.Reference number: A90759; MUID:80001710; PMID:89900
A.Contents: annotation; bilirubin binding
C.Comment: AFP is synthesized by the fetal liver and yolk sac. It occurs in the plasma c
o trace amounts after birth. The serum level in adults is usually less than 40 ng/ml. AF
C.Comment: Human AFP binds copper, nickel, and fatty acids as well as, and the bilirubin
properties.
C.Genetics:
A.Gene: GDB:AFP
A.Cross-references: GDB:119660; OMIM:104150
A.Map position: 4q11-4q13
A.Intons: 29/1, 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551
C.Superfamily: serum albumin; serum albumin repeat homology; metal binding; plasma
C.Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F.1-18/Domain: signal sequence #status predicted <SIG>
F.19-609/Product: alpha-fetoprotein #status experimental <MAT>
F.221-394/Domain: serum albumin repeat homology <SA1>
F.413-592/Domain: serum albumin repeat homology <SA2>
F.221-394/Domain: serum albumin repeat homology <SA3>
F.222/Binding site: copper (His) #status experimental
F.99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-472
F.249/Binding site: bilirubin (Lys) #status predicted
F.251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.3%; Score 1249.5; DB 1; Length 609;
Best Local Similarity 39.9%; Pred. No. 3.5e-75;
Matches 233; Conservative 116; Mismatches 231; Indels 7; Gaps 3;

3 HKSE-----VAHREKDLGEENKALVLLAFAYLOQCEPDHVKLVNEVEFAKTCVAD 57
DB HNEEGIASILDYSYCTAIEISLADLATIFPAQVGEATYKESKWKALTAIEKPTGDE 81
QY SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLOHKNDP-NLPLV 116
DB OSAGLENDLPALFLELCEHEKEILEKYG-SPDCSGSEGRNCFLAHKKPTPASIPLEQ 140
QY RPVDVMTAFHNDNEFLKRYLYEIARRHRYFAPPELLFPKRYKAAFTGCCQAADKAA 176

DB VEPVTSCEAVEDEDETFMKNFIETARRHPELVAPITLLMARYDKIIPSCKAEMAVE 200
QY CLPPLDELROGKASSAKORLKASLOKFGERRAKAMAVARSOPRAEFAEVSKITVT 236
DB CQOTAAVYTKRLRESSLLNGACAVMKNFGRTRQALITVTKLSOKFTKVNTELOKVL 260
QY DLTKVTECCGDLLECADRADLAKYICENODSISKICEKCEPDLLEKSHCIAEVND 296
DB DVAVHEHCRCRDVLDLCODGEEKIMSYICQDDTLNKTTECKLTTLERGCITHAEND 320
QY EHPADLPISLADFVSKDYCKRYAKAVYFLGMFLIXEARHDPDVSVLLPLATYETV 356
DB EPEGLSPNLNLFCDRPNQPSGSEKFLASFVHERRRHPOLAVSVILVAVAGYDEL 380
QY LEKCCAAADPHCEYKAVDEFKPLVEEPONLIKONCEFEOLGEYKFNALLVRYTKVP 416
DB LEKCPOTENPLECOGKEELQYIOESGALAKRSCGLFQKGEVYLQNAFLVATTKAP 440
QY OVSTPLVEVSNLKGKSCCKHPKAPCAEDYLSVYLNQLCVLHEKTPVSDRVTKC 476
DB QLTSELMATFRKMATATATCCQSLSEDKILACGEAADIITIGHLIRHEMPVNGVQC 500
QY CHESLVNRRPCSALEVEDETVYKPEFNATPEFHADICITLSEKROIKOTALVEYHK 536
DB CTSSTANRRPCSSLVDETVYPAFSDCKFTFHKDLQAGVALQTMKOEFLIMVKOK 560
QY PKATREOLKAVMDPEAFVEKCKADDETCEPGEKGLVAASQALGL 585
DB PRTEDQLEAVIADPSGLKECCQGEQVEYCAEFGKIKSTRALGV 609

RESULT 14

FPGO

alpha-fetoprotein precursor - gorilla

C.Species: gorilla (gorilla)

C.Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 22-Jun-1999

C.Accession: A37970

R.Ryan, S.C.; Zielinski, R.; Dugalczyk, A.

A.Title: Structure of the gorilla alpha-fetoprotein gene and the divergence of primat

A.Reference number: A37970; MUID:91169517; PMID:1706310

A.Accession: A37970

A.Molecule type: DNA

A.Residues: 1-609 <RYA>

A.Cross-references: GB:M38272; MID:9817963; PIDN:AAA73520.1; PID:9177041

C.Genetics:

A.Map position: 4q11-12

A.Intons: 29/1, 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3;

C.Superfamily: serum albumin; serum albumin repeat homology

C.Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma

F.1-18/Domain: signal sequence #status predicted <SIG>

F.19-609/Product: alpha-fetoprotein #status experimental <MAT>

F.221-394/Domain: serum albumin repeat homology <SA1>

F.413-592/Domain: serum albumin repeat homology <SA2>

F.222/Binding site: copper (His) #status predicted

F.99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-

F.249/Binding site: bilirubin (Lys) #status predicted

F.251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.0%; Score 1242.5; DB 1; Length 609;
Best Local Similarity 39.6%; Pred. No. 1e-74;
Matches 233; Conservative 117; Mismatches 232; Indels 7; Gaps 3;

3 HKSE-----VAHREKDLGEENKALVLLAFAYLOQCEPDHVKLVNEVEFAKTCVAD 57
DB HNEEGIASILDYSYCTAIEISLADLATIFPAQVGEATYKESKWKALTAIEKPTGDE 81
QY SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLOHKNDP-NLPLV 116
DB OSAGLENDLPALFLELCEHEKEILEKYG-SPDCSGSEGRNCFLAHKKPTPASIPLEQ 140
QY RPVDVMTAFHNDNEFLKRYLYEIARRHRYFAPPELLFPKRYKAAFTGCCQAADKAA 176


```

Db      141 VEPATVSCAEYEDREFNMKFIYETIARHPFLYAPTILIMARRYOKIIPSCCKAENAVE 200
Qy      177 CLIPKDELDEKASSAKORLKCASLOKGEBAFAMAVARLSQRPKAEFEVSKLYT 236
Db      201 CFQTKATVYKELRESLLNQHCAVMKNGKTRTFQAIYVTKLSOKFTKYVFTIOKLYL 260
Qy      237 DLKRVTECCGHDLLCEADRADLAVYICENODSISKLKECCCKPFLKSHCIAEVED 296
Db      261 DVANVHEHCCRGVLDLQDGEKIMYICQODTLNKTKECKKLTTLLEGQCIHAEND 320
Qy      297 EMPADPLSLADEVESKDYCKNYAEKADVFLGMELEYARRHDPYSVLLRLAKYETT 356
Db      321 EKEGELSPNLRFGDRDNQFSSGKKNIFLASFVHEYSRRHPQIYAVILRAKGVQEL 380
Qy      357 LKCCAAADPHCEYAVFDEFKPLVEBPONLIRKONCELPQDCEYKFNALVRYTKVP 416
Db      381 LEKCFQTEMLPEODDGEELQRYIQESQALAKRSGLPQKLEYYLQNFVLAAYTKAP 440
Qy      417 QVSTPLVEYSRNLGKVGSKCKHPKAPKAPCAEDLYSVNLQVLEHKTVPYSDRYTKC 476
Db      441 QLTSSSELMATRKMAATATACCOLSEDKLACGEAADITIGHLCIRHEMTVPNPGVQC 500
Qy      477 CTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQOTALVELYKHK 536
Db      501 CTSSVANNRRPCFSSLVVDETYVPARSDDKFTFHKDLQCAQVLAQTMQKQEPILNLYKOK 560
Qy      537 PKATKQQLKAVMDFAAFVCKCKKADDEKTCFAEBSGKKLVASQALGL 585
Db      561 PQTTEBQLETVIADFSGLEKCCQGOGEVCFABEGOKLISKRTALGV 609

```

RESULT 15

ABXL72

74k albumin precursor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: B41682; S02693; A05288
 R:Koskitts, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
 Mol. Endocrinol. 3, 464-473, 1989
 A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic acid during development.
 A:Reference number: A41682; MID:89313788; PMID:2747653
 A:Accession: B41682
 A:Molecule type: mRNA
 A:Residues: 3607 <MOS>
 A:Cross-references: GB:M21442; NID:g213930; PID:AAA49637.1; PID:g213931
 R:Schorpp, M.; Doebeiling, U.; Wagner, U.; Ryffel, G.U.
 J. Mol. Biol. 199, 83-93, 1988
 A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Deleted reference number: S02692; MID:88172470; PMID:2451026
 A:Accession: S02693
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-48 <SCH>
 A:Cross-references: EMBL:Z26826
 R:Molffe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, J.
 Eur. J. Biochem. 146, 489-496, 1985
 A:Title: Deinduction of transcription of Xenopus 74-kDa albumin genes and destabilization
 A:Reference number: A05288; MID:85126974; PMID:3971963
 A:Accession: A05288
 A:Molecule type: mRNA
 A:Residues: 459-502, 'L', 504-557 <MOL>
 A:Cross-references: GB:M28276
 A:Note: The authors translated the codon TAT for residue 63 as Thr
 C:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper, mones (weak bonds with these hormones promote their transfer across the membranes), thy C:Genetics:
 A:Introns: 27/1
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status predicted <PRO>

F:25-607/Product: 74k serum albumin #status predicted <MAT>
 F:33-201/Domain: serum albumin repeat homology <SA1>
 F:220-393/Domain: serum albumin repeat homology <SA2>
 F:412-591/Domain: serum albumin repeat homology <SA3>
 F:30/Binding site: copper (His) #status predicted
 F:80-88,101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-3
 F:256/Binding site: carbohydrate (Asn) (covalent) #status predicted

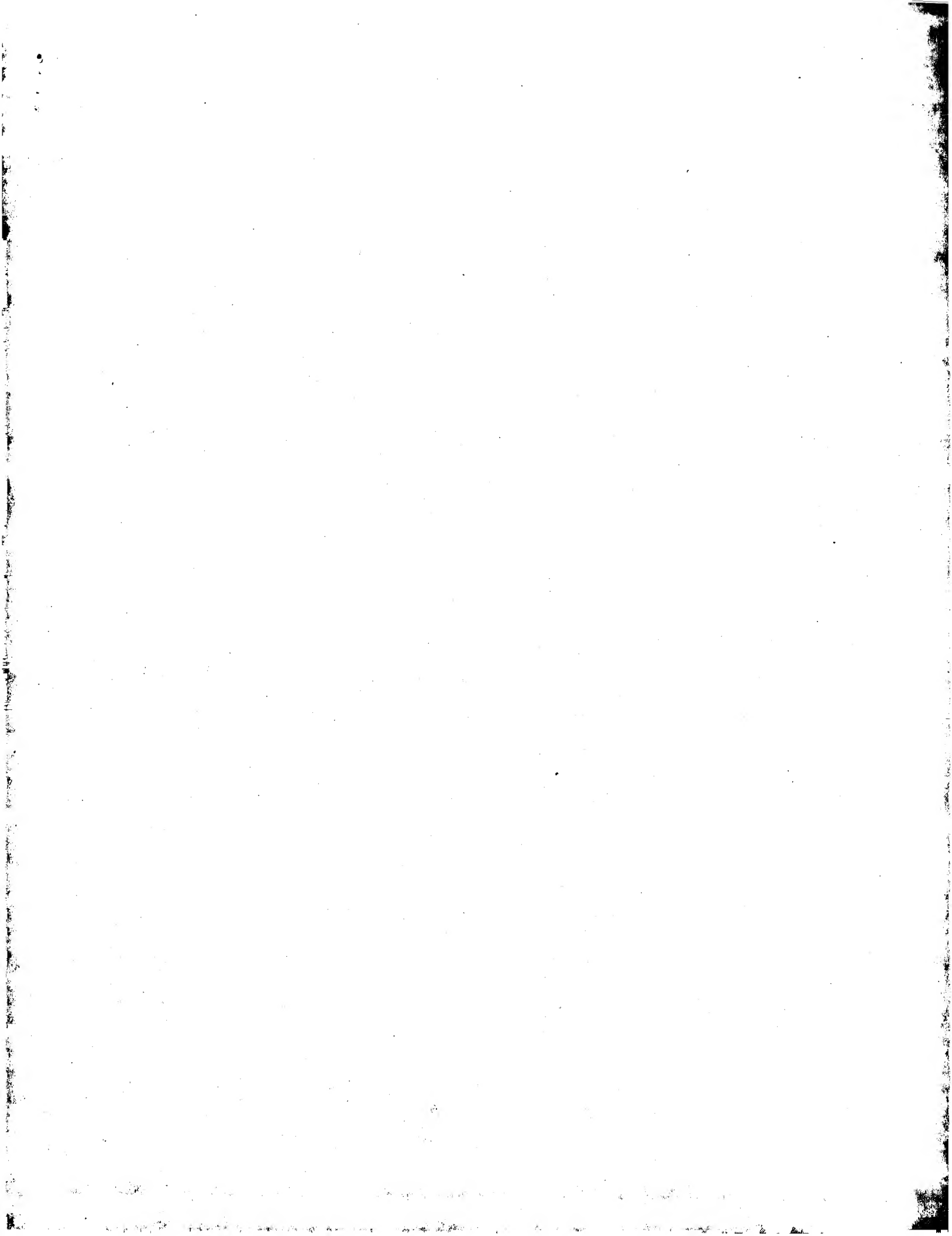
Query Match 38.88; Score 1205; DB 1; Length 607;
 Best Local Similarly 39.38; Pred. No. 3e-72;
 Matches 227; Conservative 108; Mismatches 239; Indels 4; Gaps 2;

```

Qy      3 HKSEVAFHFKDLGEENPKALVLAFAOYLQOCFEDHVKLVNTEPAKTVADSEANC 62
Db      30 HHHKADVYIALERFTFGKLTLLIVSONLQCSLEBSKLVNTEINDFASCIINDKTPC 88
Qy      63 DKSILHTLFGDKLTVALTRETGYEMADCCAKOBERNECFLOKHKDNPILRLVREVDV 122
Db      89 EKPVGTLFEDKLCADPAVGAVNYEMSKCCAKODPERAQCFKAHDEHT---SIKPEEE 145
Qy      123 MCTAFHDNETFLAKYIETARRHPFYAPABELLFPKRYKAATTECCQAADKAACILPKL 182
Db      146 TCKLKEHPDILSAFIEHARHPDLVPPAVLATLQYHKLAHCCCEEDKCKECSK 205
Qy      183 DELREDEKASSAKORLKCASLOKGEBAFAMAVARLSQRPKAEFEVSKLYTDLTKVH 242
Db      206 KQLMKOSHSLIEDKQHNFCWILDNFPEKYLKALMLAKVSHYPAERKLANHTFEVTHFI 265
Qy      243 TECCHDLEECADRADLAKYICENODSISKLKECCCKPFLKSHCIAEVEDENPADL 302
Db      266 KDCCHDMFECMTERLELFTHTQKHDELSSKLEKCNPLRLRTCYIVTLENDVPAEL 325
Qy      303 PSIAADPVESKDYCKNTAEKADVFLGMELEYARRHDPYSVLLRLAKYETTLKCCA 362
Db      326 SQPTTEFTEDPHCEYAVFDEFKPLVEBPONLIRKONCELPQDCEYKFNALVRYTKVPYSTP 385
Qy      363 AADPHCEYAVFDEFKPLVEBPONLIRKONCELPQDCEYKFNALVRYTKVPYSTP 422
Db      386 TDNPPECKYKAGADRFNENAEKERAAYLKONCDILHEHGEYLFENELLIRYTKKMPQVSDET 445
Qy      423 LVEYSRNLGKVGSKCKHPKAPKAPCAEDLYSVNLQVLEHKTVPYSDRYTKCSTELV 482
Db      446 LIGIANQMAIDIGHCACVPEENQMPACBGDLTILIGMCRQKTFINNVAACIDSY 505
Qy      483 NRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKQOTALVELYKPRATKE 542
Db      506 GMRSCFTALGPDDDYVPPVTDTPFHDDKICTANBEKOHKOKFLVLIKVSPLKKN 565
Qy      543 QLKAVMDFAAFVCKCKKADDEKTCFAEBSGKKLVASQ 580
Db      566 HIDECSAEFLKMYQKCTADEHOPCFDEKPVLIHCO 603

```

Search completed: July 22, 2003, 11:45:20
 Job time : 25 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 22, 2003, 11:38:03 ; Search time 15 Seconds

(without alignments)
1617.577 Million cell updates/sec

Title: US-09-833-118-18

Sequence: 1 DAHKSEVHAHRFKDLEGNFK.....TCFAERCKKLVAASQALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	609	1	ALBU_HUMAN
2	2942	94.8	600	1	P02768 homo sapien
3	2620	84.4	608	1	Q28522 macaca mula
4	2562	82.6	608	1	P49064 felis silve
5	2475.5	79.8	607	1	ALBU_CANFA
6	2450.5	79.0	607	1	ALBU_HORSE
7	2446	78.8	608	1	ALBU_BOVIN
8	2432.5	78.4	607	1	ALBU_SHEEP
9	2426	78.2	608	1	ALBU_RAT
10	2411.5	77.7	605	1	ALBU_PIG
11	2387	76.9	609	1	ALBU_MERIN
12	2378	76.6	608	1	ALBU_MOUSE
13	1557.5	50.2	615	1	ALBU_CHICK
14	1253.5	40.4	609	1	P19121 gallus gall
15	1249.5	40.3	609	1	P28789 pan troglod
16	1242.5	40.0	609	1	P02771 homo sapien
17	1205	38.8	607	1	P28050 gorilla gor
18	1200	38.7	609	1	P14872 xenopus lae
19	1164.5	37.5	606	1	P49066 equus cabal
20	1084	34.9	605	1	P08759 xenopus lae
21	1067	34.4	611	1	P02772 mus musculu
22	1055	34.0	599	1	P43652 homo sapien
23	944	30.4	611	1	P36923 mus musculu
24	928	29.9	608	1	P36933 rattus norv
25	747.5	24.1	608	1	P21848 salmo salar
26	742.5	23.9	608	1	Q03166 salmo salar
27	699	22.5	382	1	P21847 rana catesb
28	440.5	14.2	1423	1	Q91274 petromyzon
29	386	12.4	474	1	P02774 homo sapien
30	381	12.3	476	1	P04276 rattus norv
31	378	12.2	476	1	P53789 oryctolagus
32	372	12.0	472	1	P21614 mus musculu
33	133.5	4.3	1391	1	Q08696 drosophila

34	132.5	4.3	2230	1	G0G4_HUMAN	013439 homo sapien
35	129.5	4.2	1972	1	MYHB_HUMAN	P35749 homo sapien
36	129	4.2	3210	1	CENF_HUMAN	P49454 homo sapien
37	128	4.1	1005	1	RA50_MERJA	058718 methanococ
38	126.5	4.1	1972	1	MYHB_RABIT	P35748 oryctolagus
39	126	4.1	1189	1	SCIL_CHICK	090988 gallus gall
40	125	4.0	3038	1	TRIO_HUMAN	075962 homo sapien
41	124.5	4.0	1790	1	USOI_YEAST	P25386 saccharomyc
42	124	4.0	2653	1	CENE_HUMAN	002224 homo sapien
43	123	4.0	1057	1	EG5_HUMAN	P52732 homo sapien
44	123	4.0	1940	1	MYH3_RAT	P12847 rattus norv
45	122.5	3.9	344	1	MST1_DROHY	Q08695 drosophila

ALIGNMENTS

RESULT 1
ALBU_HUMAN STANDARD; PRT; 609 AA.
AC P02768; Q13140; Q9UTZ0; Q9UHS3; Q9P1I7; Q9P157; Q9S574;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=6196112; PubMed=3009475;
RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
Beattie W.G., Dugaiczky A.;
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within q11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=82081882; PubMed=6171778;
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
Nejarian R.C., Seeburg P.H., Wion K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli.";
RL Nucleic Acids Res. 9:6103-6114(1981).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=82105994; PubMed=6275391;
RA Dugaiczky A., Law S.W., Dennison O.E.;
RT "Nucleotide sequence and the encoded amino acids of human serum
RT albumin mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RN [4]
RP SEQUENCE FROM N.A.
PC TISSUE=Liver;
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
PC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
He F.;
RT "Functional prediction of the coding sequences of 79 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";
RN Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravsek V., Kostka V.;
RT "Complete amino acid sequence of human serum albumin.";
RN FEBS Lett. 58:134-137(1975).
RN [7]

RP SEQUENCE OF 25-609.
 RA Brown J.R., Shockey P., Behrens P.Q.;
 RL (In) Bing D.H. (eds.);
 RL The chemistry and physiology of the human plasma proteins, pp.23-40,
 RL Pergamon Press, New York (1979).
 RN
 RN SEQUENCE OF 1-455 FROM N.A.
 RP TISSUE=Liver;
 RC
 RA Menaya J., Parrilla R., Ayuso M.S.;
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
 RN
 RN PARTIAL SEQUENCE FROM N.A. (PRO2619/PRO2044/PRO1708/PRO2675).
 RC
 RA TISSUE=Fetal liver;
 RC
 RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
 RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
 RT "Functional prediction of the coding sequences of 121 new genes
 RT deduced by analysis of cDNA clones from human fetal liver.";
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
 RN
 RN SEQUENCE OF 1-26 FROM N.A.
 RA MEDLINE=86140099; PubMed=2419329;
 RC
 RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
 RT "The human albumin gene. Characterization of the 5' and 3' flanking
 RT regions and the polymorphic gene transcripts.";
 RL J. Biol. Chem. 261:3244-3251(1986).
 RN
 RN SEQUENCE OF 222-229.
 RP MEDLINE=76257808; PubMed=955075;
 RA Walker J.E.;
 RT "Lysine residue 199 of human serum albumin is modified by
 RT acetyl-salicylic acid.";
 RL FEBS Lett. 66:173-175(1976).
 RN
 RN SEQUENCE OF 25-44 AND 480-499.
 RP TISSUE=Heart;
 RC
 RA MEDLINE=95203287; PubMed=7895732;
 RC
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
 RT "The human myocardial two-dimensional gel protein database: update
 RT 1994.";
 RL Electrophoresis 15:1459-1465(1994).
 RN
 RN SEQUENCE OF 409-609 FROM N.A.
 RP TISSUE=Skeletal muscle;
 RC
 RA Strausberg R.;
 RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
 RN
 RN DISULFIDE BONDS.
 RP
 RA Sabat M.A., Stockbauer P., Moravsek L., Meloun B.;
 RT "Disulfide bonds in human serum albumin.";
 RL Collect. Czech. Chem. Commun. 42:564-579(1977).
 RN
 RN BILIRUBIN-BINDING SITE.
 RP
 RA MEDLINE=78186630; PubMed=656055;
 RC
 RA Jacobsen C.;
 RT "Lysine residue 240 of human serum albumin is involved in high-
 RT affinity binding of bilirubin.";
 RL Biochem. J. 171:453-459(1978).
 RN
 RN VARIANT CANTERBURY.
 RP MEDLINE=87157744; PubMed=3828358;
 RA Brennan S.O., Herbert P.;
 RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
 RT domain of serum albumin.";
 RL Biochim. Biophys. Acta 912:191-197(1987).
 RN
 RN VARIANTS NAG-2 AND NAG-3.
 RP MEDLINE=88068523; PubMed=3479777;
 RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
 RA Satoh C., Neel J.V.;
 RT "Amino acid substitutions in inherited albumin variants from
 RT Amerindian and Japanese populations.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
 RN
 RN

RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
 RX MEDLINE=89345611; PubMed=2762316;
 RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M.,
 RA Neel J.V., Sakurabayashi I., Putnam F.W.;
 RT "Point substitutions in Japanese allolbumins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
 RN
 RN VARIANTS MANAUS; OSAKA; NAGOTA; FUKUOKA; HONOLULU AND NEW-GUINEA.
 RP MEDLINE=90115905; PubMed=2404284;
 RA Arai K., Madison J., Shimizu A., Putnam F.W.;
 RT "Point substitutions in albumin genetic variants from Asia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
 RN
 RN VARIANT REDHILL.
 RP MEDLINE=90115852; PubMed=2104980;
 RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
 RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
 RT human serum albumin whose precursor has an aberrant signal peptidase
 RT cleavage site.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
 RN
 RN VARIANTS TORINO; VARESE AND VIBO VALENTIA.
 RP MEDLINE=91062352; PubMed=2247440;
 RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
 RA Watkins S., Putnam F.W.;
 RT "Mutations in genetic variants of human serum albumin found in
 RT Italy.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
 RN
 RN VARIANT VENEZIA.
 RP MEDLINE=91296740; PubMed=2068071;
 RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
 RA Minchiotti L., Putnam F.W.;
 RT "A donor splice mutation and a single-base deletion produce two
 RT carboxyl-terminal variants of human serum albumin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
 RN
 RN VARIANTS IOWA CITY AND KOMAGOME.
 RP MEDLINE=92052189; PubMed=1946412;
 RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
 RA Matsuda Y., I., Amaki I., Putnam F.W.;
 RT "Genetic variants of serum albumin in Americans and Japanese.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
 RN
 RN VARIANT CASEBROOK.
 RP MEDLINE=91316157; PubMed=1859851;
 RA Peach R.J., Brennan S.O.;
 RT "Structural characterization of a glycoprotein variant of human serum
 RT albumin: albumin Casebrook (494 Asp-->Asn).";
 RL Biochim. Biophys. Acta 1097:45-54(1991).
 RN
 RN VARIANTS SONDRIO AND PARIS-2.
 RP MEDLINE=92190239; PubMed=1347703;
 RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
 RA Rochu D., Porta F.;
 RT "Two allolbumins with identical electrophoretic mobility are produced
 RT by differently charged amino acid substitutions.";
 RL Biochim. Biophys. Acta 1119:232-238(1992).
 RN
 RN VARIANTS MALMO.
 RP MEDLINE=92390419; PubMed=1518850;
 RA Carlsson J., Sakamoto Y., Laurell C.-B., Madison J., Watkins S.,
 RA Putnam F.W.;
 RT "Allolbuminemia in Sweden: structural study and phenotypic
 RT distribution of nine albumin variants.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:8225-8229(1992).
 RN
 RN VARIANT HERBORN.
 RP MEDLINE=93292504; PubMed=8513793;
 RA Minchiotti L., Galliano M., Zapponi M.C., Tenni R.;
 RT "The structural characterization and bilirubin-binding properties of
 RT albumin Herborn, a [Lys240-->Glu] albumin mutant.";
 RL Eur. J. Biochem. 214:437-444(1993).
 RN
 RN

Query Match 100.0%; Score 3103; DB 1; Length 609;
 Best Local Similarity 100.0%; Pred. No. 2,7e-195;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREFKDLGSENFKALVLAFAOYLQCCPEDEHVKLVNEVTEFAKTCVADESAB 60
 DB 25 DAHSEVAHREFKDLGSENFKALVLAFAOYLQCCPEDEHVKLVNEVTEFAKTCVADESAB 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEBNECFLOHKDDPNPLPLVAPREV 120
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEBNECFLOHKDDPNPLPLVAPREV 144
 QY 121 DVMCTAFHNEEFELKLYETARHPDYAPPELLFFAKRYKAFFECOAADKAACLP 180
 DB 145 DVMCTAFHNEEFELKLYETARHPDYAPPELLFFAKRYKAFFECOAADKAACLP 204
 QY 181 KIDELDEGKASSAKORLKCASLOKGFERAFKAMAVARLSOREPKAFPAEVSCLVTDLRK 240
 DB 205 KIDELDEGKASSAKORLKCASLOKGFERAFKAMAVARLSOREPKAFPAEVSCLVTDLRK 264
 QY 241 VHTECCHGDLLECADDADLAKYICENODSISKLECCCEKPLEKSHCIAEYENDMPA 300
 DB 265 VHTECCHGDLLECADDADLAKYICENODSISKLECCCEKPLEKSHCIAEYENDMPA 324
 QY 301 DLPSLADEVESKDYCKNYAEAKDYFLGFLYEYARHPDYAVLLRLAKTYETLEKC 360
 DB 325 DLPSLADEVESKDYCKNYAEAKDYFLGFLYEYARHPDYAVLLRLAKTYETLEKC 384
 QY 361 CAADDEHECYAKVFEDEKPLVEBPONLIKONCELFOLGEYKFNALLVRYTKKVPQVST 420
 DB 385 CAADDEHECYAKVFEDEKPLVEBPONLIKONCELFOLGEYKFNALLVRYTKKVPQVST 444
 QY 421 PTLVEVSRLGKVGSKCKHPREKRMPCADYLSVNLQCVLHETPPSDRYTKCTES 480
 DB 445 PTLVEVSRLGKVGSKCKHPREKRMPCADYLSVNLQCVLHETPPSDRYTKCTES 504
 QY 481 LVNRRPCFSALEDETYVPEKFNATFTFHADICTSEKERQIKKOTALVELKHKPKAT 540
 DB 505 LVNRRPCFSALEDETYVPEKFNATFTFHADICTSEKERQIKKOTALVELKHKPKAT 564
 QY 541 KEOQLKAVMDFAAFVEKCKKADDEKTCFAEBGKKLVAAASQALGL 585
 DB 565 KEOQLKAVMDFAAFVEKCKKADDEKTCFAEBGKKLVAAASQALGL 609

RESULT 2
 ALBU_MACMU STANDARD; PRT; 600 AA.

AC Q28532;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Serum albumin precursor (Fragment).
 GN ALB.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9544;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93211971; PubMed=8460152;
 RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
 RA Dwyer J., Putnam F.W.;
 RT "cDNA and protein sequence of polymorphic macaque albumins that differ
 RT in b1/b2/b3 binding";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
 CC -1- FUNCTION: SERUM ALBUMIN. THE MAIN PROTEIN OF PLASMA. HAS A GOOD
 CC BINDING CAPACITY FOR WATER, CA++, NA+, K+, FATTY ACIDS, HORMONES,
 CC BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
 CC COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
 CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M90463; AAA36906.1; -
 DR HSSP: P02768; 1E7B.
 DR InterPro: IPR000264; Serum albumin.
 DR Pfam: PF00273; transport_prot; 3.
 DR ProDom: PD002486; Serum albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 3.
 KW Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal;
 KW Copper.
 FT NON_TER 1 1
 FT SIGNAL <1 10
 FT PROPEP 11 16
 FT CHAIN 17 600
 FT DOMAIN 17 197
 FT DOMAIN 204 389
 FT DOMAIN 396 587
 FT METAL 19 19
 FT BINDING 256 256
 FT DISULFID 69 78
 FT DISULFID 91 107
 FT DISULFID 106 117
 FT DISULFID 140 185
 FT DISULFID 184 193
 FT DISULFID 216 262
 FT DISULFID 261 269
 FT DISULFID 281 295
 FT DISULFID 294 305
 FT DISULFID 332 377
 FT DISULFID 376 385
 FT DISULFID 408 454
 FT DISULFID 453 464
 FT DISULFID 477 493
 FT DISULFID 492 503
 FT DISULFID 530 575
 FT DISULFID 574 583
 SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match 94.8%; Score 2942; DB 1; Length 600;
 Best Local Similarity 93.5%; Pred. No. 8.3e-185;
 Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHSEVAHREFKDLGSENFKALVLAFAOYLQCCPEDEHVKLVNEVTEFAKTCVADESAB 60
 DB 17 DAHSEVAHREFKDLGSENFKALVLAFAOYLQCCPEDEHVKLVNEVTEFAKTCVADESAB 76
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEBNECFLOHKDDPNPLPLVAPREV 120
 DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEBNECFLOHKDDPNPLPLVAPREV 136
 QY 121 DVMCTAFHNEEFELKLYETARHPDYAPPELLFFAKRYKAFFECOAADKAACLP 180
 DB 137 DVMCTAFHNEEFELKLYETARHPDYAPPELLFFAKRYKAFFECOAADKAACLP 196
 QY 181 KIDELDEGKASSAKORLKCASLOKGFERAFKAMAVARLSOREPKAFPAEVSCLVTDLRK 240
 DB 197 KIDELDEGKASSAKORLKCASLOKGFERAFKAMAVARLSOREPKAFPAEVSCLVTDLRK 256
 QY 241 VHTECCHGDLLECADDADLAKYICENODSISKLECCCEKPLEKSHCIAEYENDMPA 300
 DB 257 VHTECCHGDLLECADDADLAKYICENODSISKLECCCEKPLEKSHCIAEYENDMPA 316

OY 301 DLPSIADVEFSKDVCKNYAKADVFLGMLFYEXARRHDYSVLLRLATYETTTLEKC 360
 DB 317 DLPSIADVEFSKDVCKNYAKADVFLGMLFYEXARRHDYSVLLRLATYETTTLEKC 376
 OY 361 CAADPHHCYAKVDEFPPLVEBPONLKKONCELFEOIGEKFFONALLVRYTKKVPQVST 420
 DB 377 CAADPHHCYAKVDEFPPLVEBPONLKKONCELFEOIGEKFFONALLVRYTKKVPQVST 436
 OY 421 PTLVEVSNLKGVSCKCKHPEAKRMPCAEEDYLSVNLQCLVLEHKTPEVSDRYTKCTES 480
 DB 437 PTLVEVSNLKGVSCKCKHPEAKRMPCAEEDYLSVNLQCLVLEHKTPEVSDRYTKCTES 496
 OY 481 LVNRRPCSALEVDETYPPKFNATFTFHADICTLSEKROIKQALVELVKKPKAT 540
 DB 497 LVNRRPCSALEVDETYPPKFNATFTFHADICTLSEKROIKQALVELVKKPKAT 556
 OY 541 KEOLKAVDDFAAFVEKCKKADDEKTCFAEGSKKLVAAQAAL 583
 DB 557 KEOLKAVDDFAAFVEKCKKADDEKTCFAEGSKKLVAAQAAL 599

RESULT 3

ALBU_FELCA STANDARD: PRT: 608 AA.

AC P49064;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Serum albumin precursor (Allergen fel d 2).
 GN ALB.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID:9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96194824; PubMed-8647469; Hentges F.;
 RA Hiltger C., Grigioni F., Kohlen M.,
 RL "Sequence of the gene encoding cat (Felis domesticus) serum albumin."
 RL Gene 169:295-296(1996).
 CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
 BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES,
 BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
 COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: X84842; CA59279.1;
 DR HSSP: P02768; 1E7B.
 DR InterPro: IPR000264; Serum_albumin.
 DR Pfam: PF00273; transport_prot; 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 3.
 KW Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal;
 KW Copper; Allergen.
 FT SIGNAL 1 18
 FT PROPEP 19 24 BY SIMILARITY.
 FT CHAIP 25 608 BY SIMILARITY.
 FT DOMAIN 25 205 SERUM ALBUMIN.
 FT DOMAIN 212 397 ALBUMIN 1.
 FT DOMAIN 404 595 ALBUMIN 2.
 FT DOMAIN 404 595 ALBUMIN 3.

FT METAL 27 27 COPPER.
 FT DISULFD 77 86 BY SIMILARITY.
 FT DISULFD 99 115 BY SIMILARITY.
 FT DISULFD 114 125 BY SIMILARITY.
 FT DISULFD 148 193 BY SIMILARITY.
 FT DISULFD 192 201 BY SIMILARITY.
 FT DISULFD 224 270 BY SIMILARITY.
 FT DISULFD 269 277 BY SIMILARITY.
 FT DISULFD 289 303 BY SIMILARITY.
 FT DISULFD 302 313 BY SIMILARITY.
 FT DISULFD 340 385 BY SIMILARITY.
 FT DISULFD 384 393 BY SIMILARITY.
 FT DISULFD 416 462 BY SIMILARITY.
 FT DISULFD 461 472 BY SIMILARITY.
 FT DISULFD 485 501 BY SIMILARITY.
 FT DISULFD 500 511 BY SIMILARITY.
 FT DISULFD 538 583 BY SIMILARITY.
 FT DISULFD 582 591 BY SIMILARITY.
 SQ SEQUENCE 608 AA; 68659 MW; 07E629CAC5F60E5F CRC64;

Query Match 84.4%; Score 2620; DB 1; Length 608;
 Best Local Similarity 82.0%; Pred. No: 8.4e-164;
 Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;

OY 1 DAHSEVANHRRKDLGEENFKALVLAFAQYLOQCPFEHVKLVNDETEFAKTCVADESAG 60
 DB 25 EAHOSEIHRNRENDGEEHFRGLVAVFQYLOQCPFEHVKLVNDETEFAKTCVADESAG 84
 OY 61 NCDSIHTLFGDKICTVATLTRETYGEMADCCAKOPEBNEFLOHKDNPMLPRLVREV 120
 DB 85 NCEKSLHLLGDKICTVATLTRETYGEMADCCAKOPEBNEFLOHKDNPMLPRLVREV 144
 OY 121 DVMTAFHNDNEFTLKKYLYEIAIRHHPYFAPPELLFFAKRAAFTBCCQAADRAACILP 180
 DB 145 DAMCTAFHNDNEFTLKKYLYEIAIRHHPYFAPPELLFFAKRAAFTBCCQAADRAACILP 204
 OY 181 KLDELREDEGKASSAKORLKASLOKFGFRATKAAVAVLSORFRAEAEVSKVLTDLT 240
 DB 205 KYDALREKREVLASSAKERLKASLOKFGFRATKAAVAVLSORFRAEAEVSKVLTDLT 264
 OY 241 VHTCECHGDLLECCADDRADLKYICENDSISSKKECCPEPLKESHCIAEVENDEMPA 300
 DB 265 IHKECHGDLLECCADDRADLKYICENDSISSKKECCPEPLKESHCIAEVENDEMPA 324
 OY 301 DLPSIADVEFSKDVCKNYAKADVFLGMLFYEXARRHDYSVLLRLATYETTTLEKC 360
 DB 325 DLPSIADVEFSKDVCKNYAKADVFLGMLFYEXARRHDYSVLLRLATYETTTLEKC 384
 OY 361 CAADPHHCYAKVDEFPPLVEBPONLKKONCELFEOIGEKFFONALLVRYTKKVPQVST 420
 DB 385 CATDDPPACVYAHVDEFPPLVEBPONLKKONCELFEOIGEKFFONALLVRYTKKVPQVST 444
 OY 421 PTLVEVSNLKGVSCKCKHPEAKRMPCAEEDYLSVNLQCLVLEHKTPEVSDRYTKCTES 480
 DB 445 PTLVEVSNLKGVSCKCKHPEAKRMPCAEEDYLSVNLQCLVLEHKTPEVSDRYTKCTES 504
 OY 481 LVNRRPCSALEVDETYPPKFNATFTFHADICTLSEKROIKQALVELVKKPKAT 540
 DB 505 LVNRRPCSALEVDETYPPKFNATFTFHADICTLSEKROIKQALVELVKKPKAT 564
 OY 541 KEOLKAVDDFAAFVEKCKKADDEKTCFAEGSKKLVAAQAAL 583
 DB 565 KEOLKAVDDFAAFVEKCKKADDEKTCFAEGSKKLVAAQAAL 607

RESULT 4

ALBU_CANFA STANDARD: PRT: 608 AA.

AC P49822; 077705; Q9TSZ4;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Serum albumin precursor (Allergen Can f 3).

ID	ALBU_HORSE	STANDARD;	PRT;	607 AA.
AC	ALBU_HORSE			
AC	P35747;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Serum albumin precursor.			
GN	Alb.			
OS	Equus caballus (Horse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.			
OX	NCBI TaxID=9796;			
RM	(1)			
RP	SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).			
RC	TISSUE-LIVER:			
RX	MEDLINE=93345495; PubMed=8344282;			
RA	HO J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;			
RT	"X-ray and primary structure of horse serum albumin (Equus caballus)			
RL	at 0.27-nm resolution."			
RL	Eur. J. Biochem. 215:205-212(1993).			
CC	-1 FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD			
CC	BINDING CAPACITY FOR WATER, CA ⁺⁺ , NH ⁺ , K ⁺ , FATTY ACIDS, HORMONES,			
CC	BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE			
CC	COLLOIDAL OSMOTIC PRESSURE OF BLOOD.			
CC	-1 SUBCELLULAR LOCATION: Secreted.			
CC	-1 TISSUE SPECIFICITY: PLASMA.			
CC	-1 SIMILARITY: BELONGS TO THE ALB/APP/YDB FAMILY.			
CC	-1 SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: X74045; CAA52194.1; -			
DR	PIR: S34053; ABHOS.			
DR	HSSP: P02768; IE7B.			
DR	InterPro: IPR000264; Serum_albumin.			
DR	Pfam: PF00273; transport_prot; 3.			
DR	PRINTS: PR00802; SERUMALBUMIN.			
DR	ProDom: PD002486; Serum_albumin; 1.			
DR	SMART: SM00103; ALBUMIN; 3.			
DR	PROSITE: PS00212; ALBUMIN; 3.			
DR	Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal;			
KW	Copper.			
FT	SIGNAL	1	18	
FT	PROPEP	19	24	BY SIMILARITY.
FT	CHAIN	25	607	BY SIMILARITY.
FT	DOMAIN	25	204	SERUM ALBUMIN.
FT	DOMAIN	211	396	ALBUMIN 1.
FT	DOMAIN	403	594	ALBUMIN 2.
FT	METAL	27	27	ALBUMIN 3.
FT	DISULFID	77	86	COPPER (BY SIMILARITY).
FT	DISULFID	99	115	
FT	DISULFID	114	125	
FT	DISULFID	147	192	
FT	DISULFID	191	200	
FT	DISULFID	223	269	
FT	DISULFID	268	276	
FT	DISULFID	288	302	
FT	DISULFID	301	312	
FT	DISULFID	339	384	

Query Match	79.8%;	Score 2475.5;	DB 1;	Length 607;
Best Local Similarity	76.3%;	Pred. No. 2,28-154;		
Matches 445;	Conservative 69;	Mismatches 68;	Indels 1;	Gaps 1;
QY 1	DAHKSEVAHREKDIGEENFKALVLIAFQYIQCCPFEDHAKLVNEVTEFAATCYADESAE	60		
DB 25	DTKSEIARHPNDIGKEKHFQGLVVAESQYLDQCCPFEDHAKLVNEVTEFAAKCADESAAE	84		
QY 61	NCDKSLHTLWDGDKICTYATLRREYGEAMDDCAKQEPENNECFLOHKDQNPMLPRLVREV	120		
DB 85	NCDSLSHTLFQDKICTYATLRATYGLADLCCQKQEPENNECFLLHKDDHMLPRL-KRP	143		
QY 121	DVMCTAFHDNETFLIKYLYEIAIRHNPYFAPPELLFFAKRKAAFTTECCQAADRAACILP	180		
DB 144	DAQCAAAQVEDDKLQGLYVARHNPFFYFPELLFHAHEYKADTECCPADDKLACILP	203		
QY 181	KLDELREBGKASSAKORLKCSLQCFGERATKAAVAARLSORFPKAEPAEYSKLYTDLTK	240		
DB 204	KLDALEKRIILSKAKERLCKSSFQNGERAAKAMSVARLSORFPKADPAEYSKIYTDLTK	263		
QY 241	VHTCECHDILLECADRDADLAKYICENDDSISSKLKECCCKPILKESHCIATVEVDEMPA	300		
DB 264	VHKCECHDILLECADRDADLAKYICEHODSISGKAKACDDPRLQKSHCIAVEKEDDLP	323		
QY 301	DLPSLAADFEVSKDQKNYAAKQVFLCMFLYEAARRHPDYSVVLILAKYETTLK	360		
DB 324	DLPAAADFAEDKEKELCKHAKADYFLGFLYEYRRHPDYSVSLILAKYETTLK	383		
QY 361	CAADPHICVAKVDEFRPLVEEPPNLIKQNCSEFLQEGLKFPONALLVRYTKAPQVST	420		
DB 384	CAADPPACYKRVFQDFTPLVEEPPSLVKKNCDDLEEEVGEFONALLVRYTKAPQVST	443		
QY 421	PTLVEVSNLQVSGCKCKHPKAPKPAEDYLSVLMQCYLHKHTVSDRYTCCTES	480		
DB 444	PLIVEIGITLQVSGCKCKLPSENLALALNRCLVHKHTVSEKTYCCYDS	503		
QY 481	LVNRPCCSALEVDITYPKPENATFPFHADICTLSEKEROIKQOTALVELVKKRPAT	540		
DB 504	LAERPCCSALELDGIVPKPKATFTFPHADICTLPEDKQIKQOSALAEVKKRPAT	563		
QY 541	KEQLKAVMDDEFAAEVKECCAKADDEKTCFAEEGSKLVAAQSQAAL	583		
DB 564	KEQLKTVLGNPSAFYAKCCGREDKACPAEESPKLVASSQAAL	606		
RESULT 6				
ALBU_BOVIN STANDARD; PRT: 607 AA.				
ID	ALBU_BOVIN	STANDARD;	PRT:	607 AA.
AC	P02769;	O02787;		
DT	21-JUL-1986	(Rel. 01, Created)		
DT	01-FEB-1996	(Rel. 33, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Serum albumin precursor (Allergen Bos d 6).			
GN	ALB.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Holowachuk E.W., Stoltzenberg J.K., Reed R.G., Peters T. Jr.;			
RL	Submitted (Aug-1991) to the EMBL/Genbank/DBD databases.			
RN	[2]			

RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Barry T., Power S., Gannon F.;
 RL Submitted (Jul-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
 RL Submitted (Jul-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-32.
 RX MEDLINE=80024278; PubMed=488109;
 RA McGillyvray R.T.A., Chung D.W., Davie E.W.;
 RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-terminal sequence of prealbumin.";
 RL Eur. J. Biochem. 98:477-485(1979).
 RN [5]
 RP SEQUENCE OF 25-424 AND 429-607.
 RA Brown J.R.;
 RT "Structure of bovine serum albumin.";
 RL Fed. Proc. 34:591-591(1975).
 RN [6]
 RP REVISIONS TO 190-195.
 RA Brown J.R.;
 RT Submitted (APR-1975) to the PIR data bank.
 RN [7]
 RP SEQUENCE OF 402-433.
 RX MEDLINE=82023364; PubMed=7283978;
 RA Reed R.G., Putnam F.W., Peters T. Jr.;
 RT "Sequence of residues 400-403 of bovine serum albumin.";
 RL Biochem. J. 191:867-868(1980).
 RN [8]
 RP SEQUENCE OF 19-28.
 RX MEDLINE=77134075; PubMed=843354;
 RA Patterson J.E., Geller D.M.;
 RT "Bovine microsomal albumin: amino terminal sequence of bovine prealbumin.";
 RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
 RN [9]
 RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.
 RX MEDLINE=91083649; PubMed=2260975;
 RA Hirayama K., Akashi S., Furuya M., Fukuhara K.-I.;
 RT "Rapid confirmation and revision of the primary structure of bovine serum albumin by ESIMS and Frit-FAB LC/MS.";
 RL Biochem. Biophys. Res. Commun. 173:639-646(1990).
 RN [10]
 RP SEQUENCE OF 25-41.
 RX MEDLINE=88267456; PubMed=3389500;
 RA Hsieh J.C., Lin F.P., Tam M.F.;
 RT "Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing gel: a preparative method for isolating proteins for N-terminal microsequencing.";
 RL Anal. Biochem. 170:1-8(1988).
 RN [11]
 RP SEQUENCE OF 437-451.
 RA Vilbois F.;
 RT Submitted (AUG-1998) to the SWISS-PROT data bank.
 RN [12]
 RP DISULFIDE BONDS.
 RA Brown J.R.;
 RT "Structure of serum albumin: disulfide bridges.";
 RL Fed. Proc. 33:1389-1389(1974).
 CC -1- FUNCTION: SERUM ALBUMIN. THE MAIN PROTEIN OF PLASMA, HAS A GOOD BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES, BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- SIMILARITY: BELONGS TO THE ALB/AFP/NDB FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M73993; AAA51411.1; -
 DR EMBL: X58989; CAA41735.1; -
 DR EMBL: Y17769; CAA76847.1; -
 DR PIR: A38885; ABBOS.
 DR HSSP: P02768; 1E7B.
 DR InterPro: IPR000264; Serum_albumin.
 DR Pfam: PF00273; transport_prot; 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 3.
 KW Plasma; Albumin; Metal-binding; Lipid-binding; Repeat; Signal;
 KW Copper; Allergen.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 607
 FT DOMAIN 25 204
 FT DOMAIN 211 396
 FT DOMAIN 403 594
 FT METAL 27 27
 FT METAL 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 147 192
 FT DISULFID 191 200
 FT DISULFID 223 269
 FT DISULFID 268 276
 FT DISULFID 288 302
 FT DISULFID 301 312
 FT DISULFID 339 384
 FT DISULFID 383 392
 FT DISULFID 415 461
 FT DISULFID 460 471
 FT DISULFID 484 500
 FT DISULFID 499 510
 FT DISULFID 537 582
 FT DISULFID 581 590
 FT DISULFID 214 214
 FT CONFLICT 302 302
 FT CONFLICT 304 305
 FT CONFLICT 324 324
 FT CONFLICT 394 395
 FT CONFLICT 437 437
 FT CONFLICT 493 494
 SQ SEQUENCE 607 AA; 69293 MW; 39167DDE768585D4 CRC64;
 Query Match 79.0%; Score 2450.5; DB 1; Length 607;
 Best Local Similarity 75.8%; Pred. No. 9,4e-153;
 Matches 442; Conservative 71; Mismatches 69; Indels 1; Gaps 1;
 QY 1 DAHSEVARRPFDLGEENKALVLIAPQYIOCCPFEDHVKLYNETERAKTCVADSEAE 60
 DB 25 DTHKSEIARRPFDLGEHFKGLVLIAPSOYLOCCPFDEHVKLYNETERAKTCVADSEHA 84
 QY 61 NCDKSLHTLFGDGLCVATLRRETYGEMADCCAKQEPERNECFLOHDDNPNIPLRVRPV 120
 DB 85 GCEKSLHTLFGDGLCVASLRETYGEMADCCQEPERNECFLSHDDSPDLKLP 143
 QY 121 DWCTAFHNETFLKLYELARRHPYFAPLLEFARRYKNAFTCCQADKAACLP 180
 DB 144 NTLCDKFADEKFKWKYLYELARRHPYFAPLLELYANKYNGVFOCCQADKACLP 203
 QY 181 KIDLELDEKASSAKRLCASIQKGEARAFAMARISQRPKAEFAVSKLYVDLTK 240
 DB 204 KLETRMEKVALRCARQRCASIQKGEARAFAMARISQRPKAEFAVSKLYVDLTK 263
 QY 241 VHTECGHLLLECADRADLAKYICENQSISSKLECCERPLLEKSHCIAEVENDEMPA 300


```

Db 264 VHKECCHDILLECADRADLAKYICDNDJTSSKLECCDPRLEKSHCIAEVENDAIPE 323
Oy 301 DLPSTADAFVESKDKCKNYAEAKDVEFLGMFLYEVRRHDPDSVYLLRLAKYTEETLEKC 360
Db 324 NLPLTADFAEDKDYCKNYOEAKDAFLGSLFLEYSRRIPEVAVSVLLRLAKYTEATLEEC 383
Oy 361 CAADPHBECYAKVDEFEKPLVEEPONLIKONCELEFQEGEYKFONALLVRYTKKVPQVST 420
Db 384 CAKDDPHACYSYVEFDKLKHLVDEPQNLKONCDQEKGEYGFQVALLVRYTKVPQVST 443
Oy 421 PTLVEVSNLKGVSCKCKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTVSDRAVTKCTES 480
Db 444 PTLVEVSNLKGVSCKCKCKHPEAKRMPCTEDYLSLILNRLCVLHKEKTVSEKVTCKCTES 503
Oy 481 LVNRRPCFSALVEDETVYKPEFNATFTFHHADICLSKEKROIKQTALEVELVKHKPKAT 540
Db 504 LVNRRPCFSALVEDETVYKPEFNATFTFHHADICLSKEKROIKQTALEVELVKHKPKAT 563
Oy 541 KEQLKAVMDPAFAVEKCKKADKCTCPAEBGKILVAASQAL 583
Db 564 EEOQLTVMEFVAFYDKCCAADKRCACFAVEGPKLVSTOTAL 606

```

RESULT 7

ALBU_RABIT STANDARD: PRT: 608 AA.

```

ID ALBU_RABIT STANDARD: PRT: 608 AA.
AC P49065:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN 11
RP SEQUENCE FROM N.A.
RC Sheffield N.P., Syed S., Schuyler P.D.;
RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
CC BINDING CAPACITY FOR WATER, CA++, NA+, K+, FATTY ACIDS, HORMONES,
CC BUILDUP AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
CC COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U18344; AAB58347.1; -.
DR HSBP: P02768; 1E7B.
DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transport_prot. 3.
DR PRINTS: PR00802; SERUMALBUMIN.
DR ProDom: PD002486; Serum_albumin. 1.
DR SMART: SM00103; ALBUMIN_3.
DR PROSITE: PS00212; ALBUMIN_3.
DR Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; signal;
KW Copper.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.

```

```

FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; 68914 MW; CE5E92647AAFE9A2 CRC64;

```

Query Match Best Local Similarity 78.8%; Score 2446; DB 1; Length 608; Matches 433; Conservative 77; Mismatches 74; Indels 0; Gaps 0;

```

Oy 1 DAHKEVAAHRRKEDGEEFNKALVLAFAOYLOOCPFEHVKLVNEVEFAKTCVADESA 60
Db 25 EAHKSEIAHRRNDVGEHFIQVLTTFQYLOKCPYEHAKLVKCVTDLACVADESA 84
Oy 61 NCDKSLHTLFGDKICTVATLETYTGEMADCCAKOEPEBNECFLOHKDNPMLPRLVPEV 120
Db 85 NCDKSLHTLFGDKICTVATLETYTGEMADCCAKOEPEBNECFLOHKDNPMLPRLVPEV 144
Oy 121 DVMCTAFNDNETPLKLVYELIARRHRYFAVAPPELLFPAKRYKAATTECCQAADKACLLP 180
Db 145 DVLCKAFHDEKAFGHYLEVARRHRYFAVAPPELLFPAKRYKAATTECCQAADKACLLP 204
Oy 181 KLDELDRGKASASAKORLKASLOKFGERAFAVAVARLSORFPAEFAEYKLVTDLT 240
Db 205 KLDALEGSLSLSAODERLRKASLOKFGGRATKAMAVLSORFPAEFAEYKLVTDLT 264
Oy 241 VHTCECHDILLECADRADLAKYICENODSISKLECCDPRLEKSHCIAEVENDEMPA 300
Db 265 VHKECCHDILLECADRADLAKYICENODSISKLECCDPRLEKSHCIAEVENDEMPA 324
Oy 301 DLPSTADAFVESKDKCKNYAEAKDVEFLGMFLYEVRRHDPDSVYLLRLAKYTEETLEKC 360
Db 325 GLPVAABEVEEDKDYCKNYAEAKDVEFLGMFLYEVRRHDPDSVYLLRLAKYTEETLEKC 384
Oy 361 CAADPHBECYAKVDEFEKPLVEEPONLIKONCELEFQEGEYKFONALLVRYTKKVPQVST 420
Db 385 CATDDPHACYSYVEFDKLKHLVDEPQNLKONCDQEKGEYGFQVALLVRYTKKVPQVST 444
Oy 421 PTLVEVSNLKGVSCKCKCKHPEAKRMPCAEDYLSVNLQCVLHKEKTVSDRAVTKCTES 480
Db 445 PTLVEVSNLKGVSCKCKCKHPEAKRMPCEVEDYLSVNLRLCVLHKEKTVSEKVTCKCTES 504
Oy 481 LVNRRPCFSALVEDETVYKPEFNATFTFHHADICLSKEKROIKQTALEVELVKHKPKAT 540
Db 505 LVNRRPCFSALVEDETVYKPEFNATFTFHHADICLSKEKROIKQTALEVELVKHKPKAT 564
Oy 541 KEQLKAVMDPAFAVEKCKKADKCTCPAEBGKILVAASQAL 584
Db 565 NDQLKTVGFEFTALLDKCSADKRCACFAVEGPKLVSTOTAL 608

```

RESULT 8

ALBU_SHEEP STANDARD: PRT: 607 AA.

```

ID ALBU_SHEEP STANDARD: PRT: 607 AA.
AC P14639:
DT 01-APR-1990 (Rel. 14, Last sequence update)

```


DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Serum albumin precursor.
 OS Alb.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 NCBI_TaxID=9940;
 RN [1]
 RP SOURCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=90098888; PubMed-2602160;
 RA Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;
 RT "Nucleotide and deduced amino acid sequence of sheep serum albumin."
 RL Nucleic Acids Res. 17:10495-10495(1989).
 CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
 BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES,
 BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
 COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- SIMILARITY: BELONGS TO THE ALB/AF/VDL FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X17055; CAA34903.1; -.
 DR PIR: S06936; ABSH.
 DR HSSP: P02768; 1E7B.
 DR InterPro: IPR000264; Serum_albumin.
 DR Pfam: PF00273; transport_prot. 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin. 1.
 DR SMART: SM00103; ALBUMIN. 3.
 DR PROSITE: PS00212; ALBUMIN. 3.
 KW Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal;
 KW Copper.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT PROPEP 19 24 BY SIMILARITY.
 FT CHAIN 25 607 SERUM ALBUMIN.
 FT DOMAIN 25 204 ALBUMIN 1.
 FT DOMAIN 211 396 ALBUMIN 2.
 FT DOMAIN 403 594 ALBUMIN 3.
 FT METAL 27 27 COPPER (BY SIMILARITY).
 FT DISULFID 77 86 BY SIMILARITY.
 FT DISULFID 99 115 BY SIMILARITY.
 FT DISULFID 114 125 BY SIMILARITY.
 FT DISULFID 147 192 BY SIMILARITY.
 FT DISULFID 191 200 BY SIMILARITY.
 FT DISULFID 223 269 BY SIMILARITY.
 FT DISULFID 268 276 BY SIMILARITY.
 FT DISULFID 288 302 BY SIMILARITY.
 FT DISULFID 301 312 BY SIMILARITY.
 FT DISULFID 339 384 BY SIMILARITY.
 FT DISULFID 383 392 BY SIMILARITY.
 FT DISULFID 415 461 BY SIMILARITY.
 FT DISULFID 460 471 BY SIMILARITY.
 FT DISULFID 484 500 BY SIMILARITY.
 FT DISULFID 510 510 BY SIMILARITY.
 FT DISULFID 537 582 BY SIMILARITY.
 FT DISULFID 581 590 BY SIMILARITY.
 SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;
 Query Match 78.4%; Score 2432.5; DB 1; Length 607;
 Best Local Similarity 75.0%; Pred. No. 1,4e-151;
 Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

QY 1 DAHSEVAHRRKDLGEEENFKALVLIAPAYLOOCPFEDHVKLVNEYTEFAKTGVADSEAE 60
 DB 25 DTHKSELAHRRNDLGEENRFGVLVIAFSQYLQOCPFDENHVKLVNELTEFAKTGVADSEAE 84
 QY 61 NCDKSLHFLFGDKLCTVATLRETYGEMADCCAOEPEERNECFLOHKDDNPLRLVPEV 120
 DB 85 GCDKSLHTLFGDELCKVATLRETYGEMADCCOEPEERNECFLNHKDDSPDLKLT-KPEP 143
 QY 121 DVMCTAFHNEEFLLKYLLEYIETARRHPYFAPPELLFFAKRYKAFFCECAAKACLLP 180
 DB 144 DFLCAEFKADKEKFWGKLYLEVARRHPYFAPPELLYANKYNGVFOECCEAEKGLLP 203
 QY 181 KLDELDEGKASAKORLKCASLOKGEERFKMAVARISORPKAEFAVSKLYVDLTK 240
 DB 204 KIDAMEKRYLASSARQRLKASLOKGERALKMSVARLSQKPKKADFTVTVTIIVDLTK 263
 QY 241 VHECCGHDLLECCADRADLAKYICENODSISSEKLECCERPLEKSHCIAEYNDMPA 300
 DB 264 VHECCGHDLLECCADRADLAKYICDHQDASSKLECCCKPVLKSHCIAEYNDMPA 323
 QY 301 DLPSLADPFVESKDVCKNYAEADVPLGMFLYETARRHPYSVLLRLAKTETLLEK 360
 DB 324 NLPEPLADFADEKVECKNYOEAADVPLGSLYETARRHPYAVSVLLRLAKETLLEDC 383
 QY 361 CAADAPHECYAKYFDEKPLVEEPONLIRKONCELEFEOLEGEKFPONALLVRYTKRPOVST 420
 DB 384 CAEDHPACATYFDDKILHLYDEPONLIRKNCLEFEGHGYGNALIVYTKRPOVST 443
 QY 421 PTLVEYSRLNGYKSGCKCKHPEAKRMPCADYLSVNLQVLHEKTPVSDRYTKCTES 480
 DB 444 PTLVEISRLSGKYGTKCCAKPESERMPCTEDYLSLILNRLCLVHEKTPVSEKTYKCTES 503
 QY 481 LVNRRPCESLVEDETVYKPEFNAETFTFADICTLSEKROTKKQALVELYKHKPKAT 540
 DB 504 LVNRRPCESLTLDDETVYKPEFETFHADICTLPDEKQIKQALVELYKHKPKAT 563
 QY 541 KEOLKAVMDQFAFVEKCKADKCEFAEFGKLVASQAL 583
 DB 564 DEOLKAVMENFVAFVDCCAADKCEGCFVLEGRKLVASTQAL 606
 RESULT 9
 ALBU_RAT STANDARD; PRT; 608 AA.
 ID ALBU_RAT
 AC P02770; P11382;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Serum albumin precursor (Contains: Neurotensin-related peptide (NRP)).
 GN ALB.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=77249657; PubMed-893447.
 RA Straus A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
 RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
 piece. Analysis of the direct translation product of albumin
 messenger RNA".
 RT J. Biol. Chem. 252:6846-6855(1977).
 RN [3]
 RP SEQUENCE OF 25-222.
 RX MEDLINE=78109429; PubMed=564345;
 RA Isemura S., Ikenaka T.;
 RT "Amino acid sequences of fragments I and II obtained by cyanogen

	FT	DISULEID	461	472	BY SIMILARITY..
FT	DISULEID	485	501	BY SIMILARITY..	
FT	DISULEID	500	511	BY SIMILARITY..	
FT	DISULEID	538	583	BY SIMILARITY..	
FT	DISULEID	582	591	BY SIMILARITY..	
FT	VARIANT	262	262	BY SIMILARITY..	
FT	CONFLICT	174	174	V -> L.	
FT	SEQUENCE	608 AA;	68718 MM;	SBB497A28241IAB7 CRC64;	
Query Match	Best Local Similarity	78.28;	Score 2426;	DB 1;	Length 608;
Matches 428;	Conservative	73.48;	Pred. No. 3.7e-151;	Mismatches 73;	Indels 0;
				Gaps 0;	
QY	1	DAHKSEVAHFHKDGEENFKALVLAIAQYLQOCPEFDHVKLVNEVTEFAKTCVADSAAE	60		
DB	25	EAHKSELAHFKDGEQHFKGIVLIARISYLOKQRYEHIKILGEVVIDPAKTCVADENAE	84		
QY	61	NCDSKLTHTFGDKLCTVATIRETYGEMADCCAKOEPRNECFLOHKDNPILPLVPREV	120		
DB	85	NCDSKSHLTFGDKLCAIPKLRDNGELADCCAKOEPRNECFLOHKDNPILPFOREA	144		
QY	121	DVMTAFHDNEEFLFKYLVEIARHRYFAFAPELLFFPAKRYAKFAETCCQADAAACILP	180		
DB	145	EAHCTSGQENPSTSLGHLYHEVARHRYFAFAPELLYYAEAKNEVLTQOCSTSDKAACILP	204		
QY	181	KLDELDEGKAASSAKORLKCAASLOKCEBRAFKAANAVALRSQFPKAEFAEYSKLVTLTK	240		
DB	205	KLDVAKKALVAARQRMKCSSMORFERAFKAAVAARMSQRFPAEFAETTKLATDVTK	264		
QY	241	VHRECCGDLLECCDDRADLAKTYICENODSISSLKECCCEKPLEKSHCIAEVENDEMPA	300		
DB	265	INKECGDILLECCDDRADLAKTYICENODSISSLKECCCEKPLEKSHCIAEVENDEMPA	324		
QY	301	DEPSLADFEESKDVCKNVAEAKVLOMFLYEXARRHDPYSVLLRLAKTYETTLK	360		
DB	325	DLPSIADFEEDREKVCNVAEAKVPLGTFLYEXSRHDPYSVLLRLAKTYEATLEKC	384		
QY	361	CAADPHECTAKYVDEKPLVPEEFOONLIKONCELEFQLOGEKKPQNALIVRTKRVPOVST	420		
DB	385	CAEDDPACGTVLAIEFQPLVEEKNLVKTNCELYEKIGEYGFQNAVALVRYTOKAPOVST	444		
QY	421	PLTVEVRNLGKGVSKCKHPEAKRMCEADYLSVNLQVLHEKTPVSDRVTKCTES	480		
DB	445	PLTVEARNLGKGVSKCKHPEAKRMCEADYLSVNLQVLHEKTPVSDRVTKCTES	504		
QY	481	LVRNRPCEFALEVDETYVPKFNATETTFHADICTLSEKQIKQIALVLYVAKKRPAT	540		
DB	505	LVRNRPCEFALEVDETYVPKFNATETTFHADICTLSEKQIKQIALVLYVAKKRPAT	564		
QY	541	KEOLKAYMDPFAAFVEKCKRADKDCEFGFAEGKFLVAASQAAL	583		
DB	565	EDQLKTYMGDFQAEVDCCKCKRADKDCEFGFAEGKFLVAASQAAL	607		
RESULT 10					
ALBU_PIG	ALBU_PIG	STANDARD;	PRT;	605 AA.	
AC	P08835:	029018:			
DT	01-NOV-1988	(Rel. 09, Created)			
DT	01-NOV-1988	(Rel. 09, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Serum albumin precursor (Fragment).				
GN	ALB.				
OS	Sus scrofa (Pig).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
OX	NCBI_Taxid=9823;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Liver.				
RX	MEDLINE=69016582; PubMed=3174440;				
TA	Baldwin G.S., Weinstock J.,				
TA	"Nucleotide sequence of porcine liver albumin."				

RL Nucleic Acids Res. 16:9045-9045(1988).

CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD

CC BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES,

CC BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE

CC COLLOIDAL OSMOTIC PRESSURE OF BLOOD.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: PLASMA.

CC -1- SIMILARITY: BELONGS TO THE ALB/AF/VPB FAMILY.

CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X12422; CAA30970.1; -

DR EMBL; M36787; AAA30988.1; -

DR PIR; S01382; ABPGS.

DR HSSP; P02768; 1E7H.

DR InterPro; IPR000264; Serum_albumin.

DR Pfam; PF00273; Transport_prot; 3.

DR ProDom; PD002486; Serum_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

DR Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal;

KW Copper.

FT NON_TER 1 1

FT SIGNAL <1 16 BY SIMILARITY.

FT PROPEP 17 22 BY SIMILARITY.

FT CHAIN 23 605 SERUM ALBUMIN.

FT DOMAIN 209 202 ALBUMIN 1.

FT DOMAIN 401 592 ALBUMIN 2.

FT METAL 31 31 ALBUMIN 3.

FT DISULFID 75 84 COPPER (BY SIMILARITY).

FT DISULFID 97 113 BY SIMILARITY.

FT DISULFID 112 123 BY SIMILARITY.

FT DISULFID 145 190 BY SIMILARITY.

FT DISULFID 189 198 BY SIMILARITY.

FT DISULFID 221 267 BY SIMILARITY.

FT DISULFID 266 274 BY SIMILARITY.

FT DISULFID 286 300 BY SIMILARITY.

FT DISULFID 299 310 BY SIMILARITY.

FT DISULFID 337 382 BY SIMILARITY.

FT DISULFID 381 390 BY SIMILARITY.

FT DISULFID 413 459 BY SIMILARITY.

FT DISULFID 458 469 BY SIMILARITY.

FT DISULFID 482 498 BY SIMILARITY.

FT DISULFID 497 508 BY SIMILARITY.

FT DISULFID 535 580 BY SIMILARITY.

FT DISULFID 579 588 BY SIMILARITY.

FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).

SO SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1FEE CRC64;

Query Match 77.7%; Score 2411.5; DB 1; Length 605;

Best Local Similarity 76.0%; Pred. No. 3.3e-150;

Matches 438; Conservative 67; Mismatches 70; Indels 1; Gaps 1;

QY 1 DAHSEVAHREKDLGEENFALVLAFAQYLOQCPEDHYKLAVNEYTEFAKTCVADPSAE 60

DB 23 DYKSELAHREKDLGQYFEGVLVLAFAQYLOQCPEDHYKLAVNEYTEFAKTCVADPSAE 82

QY 61 NCDKSLHPLFGDKLCTVATLRETYGEMADCAKOEERNECFLOHNDNDNPLRYLREY 120

DB 83 NCDKSLHPLFGDKLCTVATLRETYGEMADCAKOEERNECFLOHNDNDNPLRYLREY 141

QY 121 DWVCTAFHNDHEFTLAKYELIARHPYFAPLLEFFAKYKAAPFECGCAAKACCLP 180

DB 142 VALCADPQDEOKQFWKCYELIARHPYFAPLLEFFAKYKAAPFECGCAAKACCLP 201

QY 181 KIDELDEGKASSAKORLCKASLOKFGERAFAKMAVAYRLSOREPKAEFAVSKLVTDLK 240

DB 202 KIDELDEGKASSAKORLCKASLOKFGERAFAKMAVAYRLSOREPKAEFAVSKLVTDLK 261

QY 241 VHECHGDLLECCADRADLAKYICENODSISKLECECEKPLLEKSHCIAEYNDMPA 300

DB 262 VHECHGDLLECCADRADLAKYICENODSISKLECECEKPLLEKSHCIAEYNDMPA 321

QY 301 DLPSIADAPFESDCKNVAEADVPLGMVLYEYARRHPYSVLLRLAKYETLLEK 360

DB 322 DLPSIADAPFESDCKNVAEADVPLGMVLYEYARRHPYSVLLRLAKYETLLEK 381

QY 361 CAADPHCEAKYFDEFKPLVEPOMLIKONCELPOLGEYKONMLVYTKKVOVST 420

DB 382 CAADPHCEAKYFDEFKPLVEPOMLIKONCELPOLGEYKONMLVYTKKVOVST 441

QY 421 PLVEVSRNLGKYSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYKCTES 480

DB 442 PLVEVSRNLGKYSKCKHPEAKRMPCAEDYLSVNLQCVLHEKTPVSDRYKCTES 501

QY 481 LVNRRCFSALFEDETVYKPEFNAETFRADICTSEKROIKKOTALVELYKHPKAT 540

DB 502 LVNRRCFSALFEDETVYKPEFNAETFRADICTSEKROIKKOTALVELYKHPKAT 561

QY 541 KEOLKAVMDPFAFVEKCKADDEKTCFPAEGEKKV 576

DB 562 EEOQLRVLGKFAFVOKCCAPDHEACFAVEGKRV 597

RESULT 11

ID ALBU_MERUN STANDARD; PRT; 609 AA.

AC G35090;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Serum albumin precursor.

GN ALB.

OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).

OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;

OC Meriones.

ON NCBI_Taxid=10047;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MGS IDR; TISSUE=Liver;

RX MEDLINE=9811663; PubMed=9455485;

RA Yoshida K., Seto-Ohshima A., Shinohara H.;

RT "Sequencing of cDNA encoding serum albumin and its extrahepatic

RT synthesis in the Mongolian gerbil, Meriones unguiculatus.";

RL DNA Res. 4:351-354(1997).

CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD

CC BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES,

CC BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE

CC COLLOIDAL OSMOTIC PRESSURE OF BLOOD.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: PLASMA.

CC -1- SIMILARITY: BELONGS TO THE ALB/AF/VPB FAMILY.

CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AB006197; BAA21765.1; -

DR HSSP; P02768; 1E7H.

DR InterPro; IPR000264; Serum_albumin.

DR Pfam; PF00273; Transport_prot; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR ProDom: PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 3.
 KW Plasma: Metal-binding; Lipid-binding; Albumin; Repeat; Signal;
 Copper.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT PROPE 19 24 BY SIMILARITY.
 FT CHAIN 25 609 SERUM ALBUMIN.
 FT DOMAIN 25 206 ALBUMIN 1.
 FT DOMAIN 213 398 ALBUMIN 2.
 FT DOMAIN 405 596 ALBUMIN 3.
 FT METAL 28 28 COPPER.
 FT DISULFD 78 87 BY SIMILARITY.
 FT DISULFD 100 116 BY SIMILARITY.
 FT DISULFD 115 126 BY SIMILARITY.
 FT DISULFD 149 194 BY SIMILARITY.
 FT DISULFD 193 202 BY SIMILARITY.
 FT DISULFD 225 271 BY SIMILARITY.
 FT DISULFD 270 278 BY SIMILARITY.
 FT DISULFD 290 304 BY SIMILARITY.
 FT DISULFD 303 314 BY SIMILARITY.
 FT DISULFD 341 386 BY SIMILARITY.
 FT DISULFD 385 394 BY SIMILARITY.
 FT DISULFD 417 463 BY SIMILARITY.
 FT DISULFD 462 473 BY SIMILARITY.
 FT DISULFD 486 502 BY SIMILARITY.
 FT DISULFD 501 512 BY SIMILARITY.
 FT DISULFD 539 584 BY SIMILARITY.
 FT DISULFD 583 592 BY SIMILARITY.
 SO SEQUENCE 609 AA: 68940 MW: 9C45F97F67E7F1A48 CRC64;
 Query Match 76.9%; Score 2387; DB 1; Length 609;
 Best Local Similarity 73.9%; Pred. No. 1.3e-148;
 Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;

DB 567 EQLKXWGEAFLEKCKQKEDKACFESTESPKLVASQKAL 608
 RESULT 12
 ALBU_MOUSE
 ID ALBU_MOUSE STANDARD: PRT: 608 AA.
 AC P07724; O61802;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB OR ALBI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
 RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fletschmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Blake J., Boffelli D., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamalya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Wittaker C., Wilming L.,
 RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE OF 99-516 FROM N.A.
 RX MEDLINE=88216123; PubMed=2452956;
 RA Minghetti P.P., Law S.W., Dugalczyk A.,
 RT "The rate of molecular evolution of alpha-fetoprotein approaches that
 of pseudogenes."
 RL Mol. Biol. Evol. 2:347-358(1985).
 RN [4]
 RP SEQUENCE OF 477-551 FROM N.A.
 RC STRAIN=BA1B/c;
 RX MEDLINE=90269606; PubMed=1971802;
 RA Boccardo C., Deschattre J., Meunier-Botival M.,
 RT "Empty and occupied insertion site of the truncated LINE-1 repeat
 located in the mouse serum albumin-encoding gene."
 RL Gene 88:181-186(1990).
 RN [5]
 RP SEQUENCE OF 25-44.
 RC TISSUE=Liver;
 RX MEDLINE=93162044; PubMed=1286668;
 RA Giometti C.S., Taylor J., Tollaksen S.L.,
 RT "Mouse liver protein database: a catalog of proteins detected by two-
 dimensional gel electrophoresis."
 RL Electrophoresis 13:970-991(1992).
 CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
 BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES,
 BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
 COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
 CC -1- SUBCELLULAR LOCATION: Secreted.

[illegible]

OY	181	LDLDELDPBGKSSAKSKORIKKASLQKFGFRAPKAAVAARLSQRPKAFPAEYSKLVTPDLTK	240
DB	205	KLDDVKEKALVSVYRORRKCSSMOKFGRAPKAAVAARLSQFPNADPAETTKLATDUTK	264
OY	241	VHTSCCHDLDLECCADDPRADLAKRYLCENODSISSKLKECCCKPLLEKSHCIAVENDEMPA	300
DB	265	VNKECCCHDLDLECCADDPRADLAKRYLCENODSISSKLQTCCKPPLKKAHCLSEVHDMPA	324
OY	301	DLPLSLADFEVSKDVCNKYNAEAKDVLGMFLYEYARRHPDYSVVLRLAKYETTTLEKC	360
DB	325	DLPLSLADFEVDEQCKNKYNAEAKDVLGTFLEYEYSRRHPDYSVLLRLAKKYEATLEKC	384
OY	361	CAADDPHECYAKYVDEKREPLVEEPPONLKKONCELPEDLOGEYKPPONALLRYTKKVPDYST	420
DB	385	CAEANPPACXYTVALEFOPPLVEEPPKNLVTKTDCDYLEKGEFGFNALLRYTORAPDYST	444
OY	421	PTLVEVSNLKGVSKCCCKHFEAKRMPCAEDYLSVLNQLCVLHEKTPVPSDRVRKCTGES	480
DB	445	PTLVEAARNLRGRVGTCKCTLPEDORLPCEVEDYLSAILNRKVLLEKTPVSHYTKCCSGS	504
OY	481	LVRNRPCHSALEVDETVYVPEKFEPAETTFNAHDITSLSEKROIKQOTALVELYVHKRKPRAT	540
DB	505	LVERPCHSALTVDETYPYKPKFAETTFNHSDDITCLPEKKEQIKQOTALVELYVHKRKPRAT	564
OY	541	KEQLKAYWDDPAFAVEKCKCKADDEKTCFAEGEKKLVAAQAL	583
DB	565	AEQLKTYWDDPAQFLDTCCKADKDCFTSTEGPNLVYTRCKDAL	607
RESULT 13			
ALBU_CHICK			
ID	ALBU_CHICK	STANDARD:	PRT: 615 AA.
AC	P1121:		
DT	01-NOV-1990 (Rel. 16, Created)		
DT	01-AUG-1992 (Rel. 23, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Serum albumin precursor.		
GN	ALB.		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauiria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
OX	NCBI_TaxID=9031;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver:		
RA	Cassady A.I., Salkilid C.K., Baverstock P., Wallace J.C.,		
RL	Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE OF 1-28 FROM N.A.		
RX	MEDLINE=83161037; Pubmed=6187737;		
RA	Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deeley R.G.,		
RT	"Hm5.5, noncoding and flanking regions of the avian very low density		
RT	apolipoprotein II and serum albumin genes. Homologies with the egg		
RT	white protein genes.";		
RL	J. Biol. Chem. 258:4556-4564(1983).		
RN	[3]		
RP	SEQUENCE OF 19-30.		
RX	MEDLINE=78019943; Pubmed=911327;		
RA	Rosen A.W., Geller D.M.,		
RT	"Chicken microsomal albumin: amino terminal sequence of chicken		
RT	prealbumin.";		
RL	Biochem. Biophys. Res. Commun. 78:1060-1066(1977).		
CC	-1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD		
CC	BINDING CAPACITY FOR WATER, CA++ , NA+ , K+ , FATTY ACIDS, HORMONES,		
CC	BIURIBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE		
CC	COLLOIDAL OSMOTIC PRESSURE OF BLOOD.		
CC	-1- SUBCELLULAR LOCATION: Secreted.		
CC	-1- TISSUE SPECIFICITY: PLASMA.		
CC	-1- SIMILARITY: BELONGS TO THE ALB/AFIP/VDB FAMILY.		
CC	-1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		

RESULT 15		
FETA_HUMAN		
ID	STANDARD;	PRT; 609 AA
AC	P02771;	
DT	21-JUL-1986 (Rel. 01, Created)	
DT	21-JUL-1986 (Rel. 01, Last sequence update)	

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-
 DE fetoprotein).
 OS AFP.
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83273664; PubMed=6192439;
 RL Moringa T., Sakai M., Wegmann T.G., Tamaki T.;
 RT "Primary structures of human alpha-fetoprotein and its mRNA.";
 RT Proc. Natl. Acad. Sci. U.S.A. 80:4604-4608(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87185438; PubMed=2436661;
 RL Gibbs P.E.M., Zielinski R., Boyd C., Dugalczyk A.;
 RT "Structure, polymorphism, and novel repeated DNA elements revealed b
 RT a complete sequence of the human alpha-fetoprotein gene";
 RL Biochemistry 26:1332-1343(1987).
 RN [3]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=93278385; PubMed=768442;
 RL McVey J.H., Michaelides K., Hansen L.P., Ferguson-Smith M.,
 RA Tiligman S., Krumlauf R., Tuddenham E.G.D.;
 RT "A G-->A substitution in an HNF 1 binding site in the human alpha-
 RT fetoprotein gene is associated with hereditary persistence of alpha-
 RT fetoprotein (HPAFP)";
 RL Hum. Mol. Genet. 2:379-379(1993).
 RN [4]
 RP SEQUENCE OF 429-556 FROM N.A.
 RX MEDLINE=83158778; PubMed=6187626;
 RA Beattie W.G., Dugalczyk A.;
 RT "Structure and evolution of human alpha-fetoprotein deduced from
 RT partial sequence of cloned cDNA";
 RL Gene 20:415-422(1982).
 RN [5]
 RP PARTIAL SEQUENCE OF 19-609.
 RX MEDLINE=81242409; PubMed=1709810;
 RA Pucci P., Siciliano R., Malorni A., Marino G., Tecce M.F.,
 RA Ceccarini C., Terrana B.;
 RT "Human alpha-fetoprotein primary structure: a mass spectrometric
 RT study.";
 RL Biochemistry 30:5061-5066(1991).
 RN [6]
 RP PRELIMINARY SEQUENCE OF 19-35.
 RX MEDLINE=77242506; PubMed=70228;
 RA Yachin S., Hsu R., Helmricks R.L., Miller J.B.;
 RT "Studies on human alpha-fetoprotein. Isolation and characterization
 RT of monomeric and polymeric forms and amino-terminal sequence
 RT analysis";
 RL Biochim. Biophys. Acta 493:418-428(1977).
 RN [7]
 RP PRELIMINARY SEQUENCE OF 19-38.
 RX MEDLINE=78001760; PubMed=711199;
 RA Aoyagi Y., Ikenaka T., Ichida F.;
 RT "Comparative chemical structures of human alpha-fetoproteins from
 RT fetal serum and from ascites fluid of a patient with hepatoma";
 RL Cancer Res. 37:3663-3667(1977).
 RN [8]
 RP PRELIMINARY SEQUENCE OF 19-39.
 RX MEDLINE=75018719; PubMed=4138095;
 RA Ruoslahti E., Pihko H., Vaheri A., Seppala M., Virolainen M.,
 RA Kontinen A.;
 RT "Alpha fetoprotein: structure and expression in man and inbred mouse
 RT strains under normal conditions and liver injury";
 RL Johns Hopkins Med. J. Suppl. 3:249-255(1974).
 RN [9]
 RP GENE STRUCTURE.
 RX MEDLINE=85182629; PubMed=2580830;
 RA Sakai M., Moringa T., Urano Y., Watanabe K., Wegmann T.G.,
 RA Tamaki T.;

RT "The human alpha-fetoprotein gene. Sequence organization and the 5' flanking region." J. Biol. Chem. 260:5055-5060(1985).

RT [10]

RP METAL-BINDING.

RA MEDLINE=79001617; PubMed=80265;

RT Aoyagi Y., Ikenaka T., Ichida F.;

RL "Copper(II)-binding ability of human alpha-fetoprotein." Cancer Res. 38:3483-3486(1978).

RN [11]

RP BILIRUBIN-BINDING.

RA MEDLINE=80001710; PubMed=89900;

RT Aoyagi Y., Ikenaka T., Ichida F.;

RL "Alpha-fetoprotein as a carrier protein in plasma and its bilirubin-binding ability." Cancer Res. 39:3571-3574(1979).

RN [12]

RP SULFATION.

RA MEDLINE=86042625; PubMed=2414772;

RT Liu M.C., Yu S., Sy J., Redman G.M., Lipmann F.;

RL "Tyrosine sulfation of proteins from the human hepatoma cell line HepG2." Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).

CC -1- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.

CC -1- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION TO THE MONOMERIC FORM.

CC -1- TO THE MONOMERIC FORM.

CC -1- TISSUE SPECIFICITY: Secreted.

CC -1- TISSUE SPECIFICITY: PLASMA. SYNTHESIZED BY THE FETAL LIVER AND YOLK SAC.

CC -1- DEVELOPMENTAL STAGE: OCCURS IN THE PLASMA OF FETUSES MORE THAN 4 WEEKS OLD. REACHES THE HIGHEST LEVELS DURING THE 12TH-16TH WEEK OF GESTATION, AND DROPS TO TRACE AMOUNTS AFTER BIRTH. THE SERUM LEVEL IN ADULTS IS USUALLY LESS THAN 40 NG/ML. AFP OCCURS ALSO AT HIGH LEVELS IN THE PLASMA AND ASCITIC FLUID OF ADULTS WITH HEPATOMA.

CC -1- PTA: INDEPENDENT STUDIES SUGGEST HETEROGENEITY OF THE AMINO-TERMINAL SEQUENCE OF THE NATURE PROTEIN AND OF THE CLEAVAGE SITE OF THE SIGNAL SEQUENCE.

CC -1- PTM: SULFATED.

CC -1- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.

CC -1- SIMILARITY: CONTAINS 3 ALBUMIN DOMAINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to license@sdb-sdb.ch).

CC -----

DR EMBL: M10949; AAA51674.1; -

DR EMBL: M10950; AAA51675.1; -

DR EMBL: V01514; CA24758.1; -

DR EMBL: M16110; AAB58754.1; -

DR EMBL: Z19352; CAA79592.1; -

DR PIR: A03234; FPHO.

DR PIR: A26624; A26624.

DR HSSP: P02768; 1E7B.

DR GLYCOSULEDB: P02771; -

DR Stena-ZDPAGE; P02771; -

DR GeneW; HGNC:317; AFP.

DR MIM: 104150; -

DR InterPro: IPR000264; Serum_albumin.

DR Pfam: PF00273; transport_prot; 3.

DR PRINTS: PR00802; SERUMALBUMIN.

DR ProDom: PD002486; Serum_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE: PS00212; ALBUMIN; 2.

DR Glycoprotein; Sulfation; Albumin; Plasma; Repeat; Metal-binding;

KW Copper; Nickel; Signal; Polymorphism.

FT SIGNAL 1 18

FT CHAIN 19 609 ALPHA-FETOPROTEIN.

FT	DOMAIN	20	205	ALBUMIN 1.
FT	DOMAIN	212	397	ALBUMIN 2.
FT	DOMAIN	404	595	ALBUMIN 3.
FT	METAL	22	22	COPPER AND NICKEL.
FT	DISULFID	99	114	
FT	DISULFID	113	124	
FT	DISULFID	148	193	
FT	DISULFID	192	201	
FT	DISULFID	224	270	
FT	DISULFID	269	277	
FT	DISULFID	289	303	
FT	DISULFID	302	313	
FT	DISULFID	384	393	
FT	DISULFID	416	462	
FT	DISULFID	461	472	
FT	DISULFID	485	501	
FT	DISULFID	500	511	
FT	DISULFID	538	583	
FT	DISULFID	582	591	
FT	CARBOHYD	251	251	
FT	VARIANT	570	570	
FT	SEQUENCE	609 AA; 68677 MW; 4D4E45820E1C2D4F CRC64;		
FT	Query Match	40.3%; Score 1249.5; DB 1; Length 609;		
FT	Best Local Similarity	39.9%; Pred. No. 1.9e-74;		
FT	Matches	235; Conservative 116; Mismatches 231; Indels 7; Gaps 3;		
QY	3 HKSE-----VAHREFDGEENFKALVLAFAQYLODCPEHDKVLYNEVEFAKTVAD 57			
DB	22 HNEVGIASILDSDYCTAEISLADLTTFQAFVQEAATYKEVSKVMDALPAIEKPTGDE 81			
QY	58 SAENCDKSLHLEFGDKCTVATLRRTYGMADCCAKOBERNECPLOHKDNDP-NLPRLY 116			
DB	82 OSSGCELOLPALFELELHEKEILEKIGH-SPCCSQSEBGRNCLAKKPPASIPLEQ 140			
QY	117 PREVVMCTAFHNDNEFTLKKYLYEIRARRHYFAPELLFFAKRYKAAFTGCCQADRAA 176			
DB	141 VEPYTSCEAVYEDRETFMKNFIYIARRHPLVPTLLIMARVYDKIIPSCKAENAVE 200			
QY	177 CLPRYLDELDRGSKSSAKQRLKASLOKFGRRAKAAVAVLRSFRKAEFAEVSKITV 236			
DB	201 CQTAATATYVELRRESSILNDHACAVMNFGRTOALTIVTKLSQKFTKVNFTETOKLV 260			
QY	237 DLTQVTECHCHDLECCADDRADLAKYICENODSISSKLECCERPLLEKSHCIAEVND 296			
DB	261 DVANHEHCRCRDVLDCLDQDGKIMSYICSQDITLSNKTTECCKLTLEKGCIIHAEND 320			
QY	297 EMPADLPISLADEVESKDVCKNYABAKOVFLGMLFYEARRRHDPYSVALLRLAKTYETT 356			
DB	321 EKPEGLSPNLNFFLDRODFNPFSSGCKNFIASFVHEYSRRHPOLAVSYILVAKGYOL 380			
QY	357 LEKCAADPHHCYAKVDEFRPLVEEPONLIKONCELEFEOLEKCFKFNALLVRYTKVP 416			
DB	381 LEKCFOTENPLKQGEELQKYQESQALAKKSGLFQKIGEYLYLDNALFVAITKAP 440			
QY	417 QVSTPTLVEVSNLKGKSCCKHPKAPKMPKAEEDYLSVNLQCLVLEHEKTPVSDRYTKC 476			
DB	441 QUTSSELMATITRKMAATATCCQSLSEDKLACGEBAAOIIIGHLCIRHEMPVNVGVOGC 500			
QY	477 CTESLVNRRPCFSALVDETVYVPEFNAETFTFHADICTLSEKROIKQATALVELVHK 536			
DB	501 CTSSVANRRPCSSLVVDVETVPAPFSDKFTFHKDLCAQAGVALQTMKOEPLMLVVKOK 560			
QY	537 PKATKEOLKAVMDPAFAVEKCKKADDETCGAEFGKLVASQALGL 585			
DB	561 POLTEOLEAVYADFSGLLEKCCQGOEDEVCFAEFGKLVISKTRALGV 609			

Search completed: July 22, 2003, 11:44:04

Job time : 17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 22, 2003, 11:42:28 ; Search time 39 Seconds
(without alignments)
3090.707 Million cell updates/sec

Title: US-09-833-118-18
Perfect score: 3103
Sequence: 1 DAHKESEVAHREKDLGEENFK.....TCFAEEGKKLVASQAALGL 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp Unclassified:*
15: sp_rvirus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2504	80.7	608	5	Q95VB7 schistosoma
2	1242	40.0	626	13	Q8UW05 ambystoma m
3	1087	35.0	624	13	Q8UW06 ambystoma t
4	955	30.8	603	13	Q9YGH6 rana shqipe
5	928.5	29.9	614	13	Q91134 naja naja (
6	713	23.0	417	11	Q8R0J9 mus musculu
7	376.5	12.1	484	13	Q9W6F5 gallus galli
8	373	12.0	476	11	Q9CY31 mus musculu
9	372	12.0	476	11	Q9CY31 mus musculu
10	365	11.8	551	13	Q42279 petromyzon
11	331	10.7	122	13	Q90WZ8 latius argen
12	290	9.3	123	13	Q90WZ6 poephila gu
13	264	8.5	135	11	Q63205 rattus norv
14	188	6.1	173	2	Q9JMX8 helicobacte
15	184	5.9	1819	16	Q9ZLV0 helicobacte
16	184	5.9	1927	16	Q25262 helicobacte

17	162.5	5.2	44	6	Q95MC2	Q95MC2 equus caball
18	162.5	5.2	680	5	Q9V6S8	Q9V6S8 drosophila
19	161	5.2	1079	3	Q96V11	Q96V11 pneumocysti
20	156	5.0	1026	3	Q74669	Q74669 pneumocysti
21	154.5	5.0	3843	5	Q905D0	Q905D0 drosophila
22	153.5	4.9	3843	5	Q9V094	Q9V094 drosophila
23	151.5	4.9	1605	11	Q99PL5	Q99PL5 mus musculu
24	150	4.8	1065	3	Q01828	Q01828 pneumocysti
25	149	4.8	40	6	Q9TRA5	Q9TRA5 oryctolagus
26	148	4.8	1028	3	Q74668	Q74668 pneumocysti
27	142.5	4.6	2736	10	Q9LJ60	Q9LJ60 arabidopsis
28	141.5	4.6	1560	5	Q26644	Q26644 strongyloce
29	138.5	4.5	1069	3	Q96V12	Q96V12 pneumocysti
30	138.5	4.5	1560	4	Q9P2E9	Q9P2E9 homo sapien
31	137.5	4.4	1348	16	Q8YK55	Q8YK55 anabaena sp
32	135.5	4.4	2931	5	Q9WZC6	Q9WZC6 drosophila
33	134.5	4.3	1005	5	Q01794	Q01794 pneumocysti
34	134.5	4.3	1154	5	Q9V6S9	Q9V6S9 drosophila
35	134	4.3	1083	3	Q12075	Q12075 pneumocysti
36	134	4.3	1780	4	Q9URF5	Q9URF5 homo sapien
37	132.5	4.3	1070	10	Q9SV36	Q9SV36 arabidopsis
38	131	4.2	1076	3	Q01830	Q01830 pneumocysti
39	128.5	4.1	607	5	Q9V6S7	Q9V6S7 drosophila
40	128.5	4.1	1051	3	Q01694	Q01694 pneumocysti
41	128	4.1	868	12	Q9J1A8	Q9J1A8 barley yell
42	128	4.1	969	5	Q9ND19	Q9ND19 plasmodium
43	128	4.1	1092	3	Q9UYR2	Q9UYR2 pneumocysti
44	126.5	4.1	1641	11	Q88528	Q88528 mus musculu
45	126.5	4.1	1927	5	Q25142	Q25142 halocynthia

ALIGNMENTS

RESULT 1	ID	Q95VB7	PRELIMINARY;	PRT;	608 AA.
AC	Q95VB7				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)				
DE	Albumin.				
OS	Schistosoma mansoni (blood fluke).				
OC	Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeiida;				
OX	Schistosomatidae; Schistosomatidae; Schistosoma.				
NC	NCBI_TaxID=6183;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Osman A., Asahi H., Stadecker M.J., Loverde P.T.;				
RT	"Albumin precursor homolog is a novel T helper cell immunogenic egg component in murine infection with Schistosoma mansoni."				
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF18550; AAL08579.1; -				
DR	InterPro; IPR000264; Serum_albumin.				
DR	Pfam; PF00273; transprot_prot; 3.				
DR	Prodom; PD002486; Serum_albumin; 1.				
DR	PROSITE; PS00212; ALBUMIN; UNKNOWN-2.				
SQ	SEQUENCE 608 AA; 68225 MW; ESEAB28E1C66E54 CRC64;				
Query Match	80.7%; Score 2504; DB 5; Length 608;				
Best Local Similarity	76.3%; Pred. No. 3.4e-183;				
Matches	445; Conservative 79; Mismatches 59; Indels 0; Gaps 0;				
QY	1 DAHKESEVAHREKDLGEENFKALVLAFAQYLQCCPEEDHVKLVNVEYEFKTKVADESA 60				
DB	: : : : : : : : : :				
QY	25 DAHKESEVAHREKDLGEENFKALVLAFAQYLQCCPEEDHVKLVNVEYEFKTKVADESA 84				
DB	: : : : : : : : : :				
QY	61 NCKRSHTLEFGDKLCTVATREYTGEMADCCAKOEPRNCELOHDDNPNLEPRV 120				
DB	: : : : : : : : : :				
QY	85 NCKRSHTLEFGDKLCTVATREYTGEMADCCAKOEPRNCELOHDDNPNLEPRV 144				
DB	: : : : : : : : : :				
QY	121 DVNCTAFHNDNEFLKYLVEIARRHPYFAPRLFFAKRYKAAFTCCCAADKACLLP 180				
DB	: : : : : : : : : :				

DB 145 EAMCTSFQENAVTEMGHYLHEVARRHPYFAPELLLYAEKYSALIMECCGGEADKAACITP 204
QY 181 KLDELROGKASAKORLTKCASLOKFGERAFAKAMAVARLSOREPFAEVAESKLVTDLTK 240
DB 205 KIDAIKEKALASVQORLKCSSLQRFQGRAPKAMAVAKMSQKFPADRAETIKLTDLTK 264
QY 241 VHTCECHGDLLECADDRLADLAKYICENODSISIKLECCCEKPLLEKSHCIAEVENDEMPA 300
DB 265 LTEECCHGDLLECADDRAELAKYMCENODASISSIKLOACCCKPVLKSHSLSEVENDIPA 324
QY 301 DLPSTAAOFVESKOVCKNNAEAKDVFGLMFLYEVARRHPDVSVLLLLAKTYETTLK 360
DB 325 DLPSTAAOFVEDEKCKNNAEAKDVFGLMFLYEVARRHPDVSVALLLAKTYETTLK 384
QY 361 CAADPHCEYAKVFEFFPFLVEEPONLIRKONCELEFGEQYKFNALLVRYTKKPOYST 420
DB 385 CAADPHCEYAKVFEFFPFLVEEPONLIRKONCELEFGEQYKFNALLVRYTKKPOYST 444
QY 421 PTLVEASRNGLGVGSKCKCKHPEAKRMPCAEDYLSVNLQCVLHETKTPVSDRYTKCCTES 480
DB 445 PTLVEASRNGLGVGSKCKCKHPEAKRMPCAEDYLSVNLQCVLHETKTPVSDRYTKCCTES 504
QY 481 LVNRPCFSALEVDETYVPKEFNAETPFPHADICLSKEROIKKOTLVLVELYKHKPKAT 540
DB 505 VVERPCFSALPVDDETYVPKEFNAETPFPHADICLSKEROIKKOTLVLVELYKHKPKAT 564
QY 541 KEOLKAVMDPFAEVEKCKCKADDKETCFEAEGKRLVAASQAAL 583
DB 565 GPQLRTVGEFTAFILDKCKKAEKDEKACFSEDSQKRLVASSQAAL 607

RESULT 2

Q80W05 PRELIMINARY; PRT: 626 AA.
AC 080W05:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ambystoma maculatum (spotted salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidae; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=43114;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER.
RA Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
RT Ambystoma texanum."
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF217183; AAL56646.1; -
DR InterPro: IPR00264; Serum_albumin.
DR Pfam: PF00273; transport_prot. 3.
DR PRINTS: PR00802; SEROMALBUMIN.
DR PRODOM: PD002486; Serum_albumin. 1.
DR SMART: SM00103; ALBUMIN. 3.
DR PROSITE: PS00212; ALBUMIN; UNKNOWN_2.
KW Signal.
FT SIGNAL. 1 24 POTENTIAL.
FT CHAIN 25 626 SERUM ALBUMIN.
SQ SEQUENCE 626 AA; 70677 MW; 9D66F57F174AC23F CRC64;

Query Match 40.0%; Score 1242; DB 13; Length 626;
Best Local Similarity 40.1%; Pred. No. 1e-86; Indels 6; Gaps 3;
Matches 237; Conservative 110; Mismatches 238;

QY 1 DAHKESEVARRFDL-----GEENFKALVLIAPAOYLQOCPEFDHVLVNEVTEFACTVAD 56
DB 28 EGHVNDPRLIGDILPMIGVNDKSLVLAASQMLPLCPYEHDLORVEDVYQMLADLCKAG 87
QY 57 ESAENCKDSLHTLFGDKLCTVATLREYGEADCCAKOPEPNECFLOHKDND-PNLPL 115

DB 88 ARHNAKASPMITIIIDELCKRPENAEKYPFHQECCKEDPERHKCFVHEKNAHBEILRY 147
QY 116 VREPDVWCTAFHONDEFLLKLYEIRRRHPYFAPELLLYAEKYSALIMECCGGEADKAACITP 175
DB 148 VRPADEQICDHAERKGLPILYFMLAIGHHMIIPAILGFAOFGDVISHCADVETA 207
QY 176 A-CLIPKIDELREDEKASSAKORLTKCASLOKFGERAFAKAMAVARLSOREPFAEVAESKL 234
DB 208 GOCFNDKRPKHQVELEYCALQKHNCYILDOFKERALAYAVAKMSQKFPADRAETIKLTDLTK 267
QY 235 VTDLTKVHTCECHGDLLECADDRLADLAKYICENODSISIKLECCCEKPLLEKSHCIAEVE 294
DB 268 VPDVHNLHQTCGSGMMAEMLEKMTIKICEKDELTHTHECCDKPYLESACTIITLP 327
QY 295 NDEMPADP.LSLADEVESKOVCKNNAEAKDVFGLMFLYEVARRHPDVSVALLLAKTYETTLK 354
DB 328 NDQKPADLSPKVPYHIDPEYCKLYTEGGDTFMGFLYECARRHODYSPENLLRSGSYE 387
QY 355 TTLEKCAADPHCEYAKVFEFFPFLVEEPONLIRKONCELEFGEQYKFNALLVRYTKK 414
DB 388 EFLKCCAAEGHNECLAKTEESLKEIESVTLTKTNGALDKLSYLFQNLIRKYAVR 447
QY 415 VPQSTPFLVEASRNGLGVGSKCKCKHPEAKRMPCAEDYLSVNLQCVLHETKTPVSDRYTK 474
DB 448 MPALSEOSILRTKSMITIGECCKHPEAKRMPCAEDYLSVNLQCVLHETKTPVSDRYTK 507
QY 475 KCCSTELVNRPCFSALEVDETYVPKEFNAETPFPHADICLSKEROIKKOTLVLVELYKHKPKAT 534
DB 508 OCCSHLSSTQTPCFALPYDETYVPPLSVASFNNDELCTTSPEDOSKQOVLIRLMK 567
QY 535 HKPKATKEOLKAVMDPFAEVEKCKCKADDKETCFEAEGKRLVAASQAAL 583
DB 568 QYRHWTDQOLTKCYVNFVPMYDQCKADNHNECFALLEGAKLIDACKAILAV 618

RESULT 3

Q80W06 PRELIMINARY; PRT: 624 AA.
AC 080W06:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ambystoma texanum (Smallmouth salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidae; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8304;
RN [1]
RP SEQUENCE FROM N.A.
RA Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
RT Ambystoma texanum."
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF217182; AAL56645.1; -
DR InterPro: IPR00264; Serum_albumin.
DR Pfam: PF00273; transport_prot. 3.
DR PRODOM: PD002486; Serum_albumin. 1.
DR SMART: SM00103; ALBUMIN. 3.
DR PROSITE: PS00212; ALBUMIN; UNKNOWN_1.
KW Signal.
FT SIGNAL. 1 24 POTENTIAL.
FT CHAIN 25 624 SERUM ALBUMIN.
SQ SEQUENCE 624 AA; 70321 MW; DE08533BF4955EF7 CRC64;

Query Match 35.0%; Score 1087; DB 13; Length 624;
Best Local Similarity 37.7%; Pred. No. 7.1e-75; Indels 12; Gaps 5;
Matches 214; Conservative 102; Mismatches 240;

QY 14 LGSENFKALVLIAPAOYLQOCPEFDHVLVNEVTEFACTVADSAENCKDSLHTLFGDK 73

```

Db 45 IGVENAKALAMALFSGMLSKCPHHEQVQRVRYNMDIADLCISGAKHGDCGKSVMTIILNE 104
QY 74 ICTVATILRETYGEMADOCACOEPRNECFLOHKDNP-NLPRLVREPVDMCTAFPHDNEE 132
Db 105 ICTPENPEKYPHHEGCCCKEDPERKCFLEHSTDPKETEYKSPEDICDHAENRD 164
QY 133 TELKKLYELIARRHPFYFABELLFFAKRYKAFTTECCQ--AADKACILPKILDELREGKA 191
Db 165 EFLGHYIHKVASSHTMYPRALISFTLHFGIYSHOCKDEATVGOCISEKMPAHKEVEH 224
QY 192 SSKAKOLKASLOKFEGRARFKAANAVARLSORFPAEFAEYSKLVTDLTKYHTECCGDL 251
Db 225 VCAVQKHNCTYILQNFERRALASKAAHACSKFPHASEENVQRLTDGIVLHQCCGDDM 284
QY 252 ECADDDADLAKYICENODSISSKLEKCEKPLEKSHCIAEVENDEMPADLPISLADVE 311
Db 285 ACAAEKMKLTQYCEK-----KKCEKPYLERSECIYRLPNDEKPALESEVRYTFD 336
QY 312 SKDCKNYAEAKDVFLGMFLYEXARRHPDYSVLLRLAKTYETLEKCCAAADPHCEYA 371
Db 337 DPEVCKRFKEGDAFNGRFLCDYAKIHPSHAEILNRIASGLEKAYKTCAGAHNCEIA 396
QY 372 KYDEKPLVEEPQNLKONCELFEOUGEYKFONALLVRYTKVPOVSTPLVEVSRNLG 431
Db 397 KEETELRHEIEAKSKTKLKTGALKEKGYHFNQIMIVRYTGILPSSDAFLYITKTLT 456
QY 432 KVGSKCKHPEAKRMPCAEDYLSVLNOLCVLHEKTPV-SDRYTKCCTESLVNRRPESA 490
Db 457 NIOKCKPLPEDQCMCSSEGGLMVFRAQIC-QOKCTPENEKLAHCKSDLSFTTPEFAA 515
QY 491 LEVDETVYVPEKFAETFTFHADICTLSEKEROIKQOTALVELVYKHPKATREOLKAMVD 550
Db 516 LVDETVYVPAVPAESPFENFDECTPSEADLOAKKQTFMLHIVRTHKIIDQYKTISEK 575
QY 551 FAAFVEKCCADKDKETCPAEKGYAA 578
Db 576 FLAMGCGCCAKADQNECFATEGAKLVEA 603

```

RESULT 4

```

QY9G6 PRELIMINARY; PRT; 603 AA.
AC QY9G6;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Rana shqiperica.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=44336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BUSHT; TISSUE-LIVER;
RA Uzzell T., Holtz H.;
RT "Albumin cDNA sequence of Rana shqiperica: evolutionary changes in
RT frog albumins." to the EMBL/GenBank/DBJ databases.
RL Submited (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40452; AAD09358.1; -.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot. 3.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KM Signal.
FT NON TER 1
FT SIGNAL <1 23 POTENTIAL.
FT CHAIN 24 603 SERUM ALBUMIN.
SQ SEQUENCE 603 AA; 69293 MW; 340D3723FA010C99 CRC64;

```

```

Query Match 30.8%; Score 955; DB 13; Length 603;
Best Local Similarity 33.6%; Pred. No. 8.3e-65;
Matches 190; Conservativity 119; Mismatches 241; Indels 16; Gaps 3;

```

```

QY 12 KDIGENFKALVLIARQYIQCCPFEDHKLAVNEVTEFAKTCYADSAENCKSLHTLREG 71
Db 37 KAVGKPAVEKLVLMVAODEKESCLDEHKLQAKIIEAVDNCCKHPEAECKKPALETLYH 96
QY 72 DKICVATILRETYGEMADOCACOEPRNECFLOHKDNP-LPRLV-REPVDMCT 125
Db 97 DIVCKEEDIDQLYPMTECCGKAERTKCFYHRE-----VREEYKIPNIESC 148
QY 126 AFHDNEETELKKLYELIARRHPFYFABELLFFAKRYKAFTTECCQADKACILPKILDEL 185
Db 149 EKHENQGRANFYSYLINAKHSHLYPRAVLGFIQYNEITTECCAADKACGEMPMY 208
QY 186 RDEKASSAKOLKASLOKFEGRARFKAANAVARLSORFPAEFAEYSKLVTDLTKYHTEC 245
Db 209 KKLTVNLEDKHKOCRYLKEFPERSQALTLVQSORFGNAKYDVEKVTIEIHLNEDC 268
QY 246 CHGDLLECADRDADLAKYICENODSISSKLEKCEKPLEKSHCIAEVENDEMPADLPISL 305
Db 269 CKGDVAECMIERMEATEHICLAKELSSKLSLDCAGVLERTPCILALPNEE--PDLP 326
QY 306 AADFVSKDCKNYAEAKDVFLGMFLYEXARRHPDYSVLLRLAKTYETLEKCCAAAD 365
Db 327 LKYEYDEHVCENYQDKRYTLHFNHDSYRSHQESSPOGCLVRSKGFEMLEKCCASAN 386
QY 366 PHECYAVDEKPLVEEPQNLKONCELFEOUGEYKFONALLVRYTKVPOVSTPLVE 425
Db 387 SAECLDAPKILBAALKENEISKONGALKEKGFNDYFOLLVRYFGKMPYQTAQTYVE 446
QY 426 VSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNOLCVLHEKTPVSDRYTKCCTESLVNRR 485
Db 447 LTRMAKIGYVCGGLPNKKQPCAEKLDLILGEMEREKRTINDVHVCVDSYANRR 506
QY 486 PCFSALVEVDTVYVPEKFAETFTFHADICTLSEKEROIKQOTALVELVYKHPKATREOLK 545
Db 507 PCFTKLGPYANVPAWPWDESKLFTADMGCSADDQKTLVLLVEFLMKKPCGCKEKL 566
QY 546 AVMDFAAFVEKCCADKDKETCPAE 571
Db 567 EVIESPFRTVVECCAENQOACFDEK 592

```

RESULT 5

```

QY91134 PRELIMINARY; PRT; 614 AA.
AC QY91134;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE Cobra serum albumin.
OS Naja naja (Indian cobra).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Elapidae; Elapinae; Naja.
OX NCBI_TaxID=35670;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER.
RX MEDLINE=96145734; PubMed=8561913;
RA Wang X., Hansen H., Haysteen B.;
RT "Evidence of the coevolution of snake toxin and its endogenous
RT antitoxin. Cloning, sequence and expression of a serum albumin cDNA of
RT the chinese cobra."
RL Biol. Chem. Hoppe-seyler 376:545-553(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RX MEDLINE=9343893; PubMed=8343135;
RA Shao J., Shen H., Havsteen B.;
RT "Purification, characterization and binding interactions of the

```


OY 1 DAHSEVAHREKDLGEENFKALVILFAOYLQOCPEFEDHVLVNEVEFEAKCYADESAE 60
 DB 22 DYEDKONCENLAMEGKDEFRSLSLIVSRKSSSFPEOVNOLVEVYSLTECCAGADP 81
 OY 61 NCDKSLHTLFEGDKICTYATLTRETYGEMADCCAKOPEPNEFLOHKDNPMLPRLVREV 120
 DB 82 TCYOTRSELSVKSCESDAPPVHGPRPECCCKEGLERKLMAALSHOPOEPYVEPTN 141
 OY 121 DVMCTAFHDNEETFLKTYELIARHHPYFVAPELLFFAKRYKAAPTECCOADAACLLP 180
 DB 142 DEICAFRRDPKGRADOPLEYSSNYGAPRLPLVATYKNTLSMGSCCTSANPLVCVK 201
 OY 181 KLDELREGKASSAKORLKASLOKFGERAFAKANAVALSORFPAEFAEYSKLVTDLTK 240
 DB 202 ERLQMKHLSTLTMSNRV-CSOYAAVGKESKRLSHLILAKQVPTANLENVLPILAEDPTE 260
 OY 241 VHTECCHDLEEC-ADDRADLAKYICENODSISKLKCCCKPFLKSHCIAEYNDM- 298
 DB 261 ILSRCESTSEDCASSELPEHTIKQNLKSKNSFECCO-----ENTPMN 307
 OY 299 -----PADLPSELADPVESEKDYC-KNYAEAKDVLGMFLYFARHHPDYSVL 345
 DB 308 IFMCTFEPAPAEPIQLP-AIKLPKGDLCGOSTTQAMD-----QTFELSKRTQVPEVFL 361
 OY 346 LILAKTYETTLKCCAAADPHECYAKVDEKPLVE-EPONLIKQNCLEFEOLGEYKF- 403
 DB 362 SKVLEPTLK-TLRECCDTQDSVAC---FSTQSPILKQNLSTIEKGEOMCADSENFT 416
 OY 404 --QNALVRYTKKYPQVSTPPLVEYSRNLGKVGSKCC 438
 DB 417 EYKKKLAERLRTKPTNTPSPALKDMVEKHSDFASKCC 453

RESULT 10

ID 042279 PRELIMINARY; PRT: 551 AA.
 AC 042279;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Serum albumin AS (Fragment).
 GN AS.
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hypercoartia;
 CC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX Nchi_TaxID=7757;
 RN [1]
 RP SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.
 RC TISSUE=LARVAL LIVER;
 RX MEDLINE=98428063; PubMed=9755481;
 RA Filosa M.F., Adam I., Robson P., Helnig J.A., Smith K., Keeley F.W.,
 Youson J.H.;
 RT "Partial clone of the gene for AS protein of the lamprey Petromyzon
 marinus, a member of the albumin supergene family whose expression is
 restricted to the larval and metamorphic phases of the life cycle.";
 RL J. Exp. Zool. 282:301-309(1998).
 RN [2]
 RP DEVELOPMENTAL STAGE.
 RX PubMed=10998585;
 RA Dauls M.H., Filosa M.F., Youson J.H.;
 RT "An albumin-like protein in the serum of non-parasitic brook lamprey
 (Lampetra appendix) is restricted to preadult phases of the life cycle
 in contrast to the parasitic species Petromyzon marinus.";
 RL Comp. Biochem. Physiol. 127B:251-260(2000)
 CC -I- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
 BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES,
 BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
 COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
 CC -I- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -I- DEVELOPMENTAL STAGE: LARVAL (AMMOCETE), METAMORPHOSING, AND
 CC JUVENILE INDIVIDUALS, BUT NOT IN SEXUALLY MATURE ADULTS.
 CC -I- DOMAIN: COMPOSED OF AT LEAST THREE HOMOLOGOUS DOMAINS.

CC -I- MISCELLANEOUS: IN THE SEA LAMPREY, THERE ARE TWO FORMS OF ALBUMIN,
 AS AND SDS-1.
 CC -I- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
 DR EMIL: AF031134; AAC63407.1; "
 DR InterPro: IPR000264; Serum_albumin.
 DR Pfam: PF00273; transport_prot. 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin; 2.
 DR SMART: SM0103; ALBUMIN; 2.
 DR PROSITE: PS00212; ALBUMIN; 2.
 KW Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Glycoprotein.
 FT NON_TER 1
 FT CHAIN 1
 FT FT <1 551 SERUM ALBUMIN AS.
 FT REPEAT <1 127 3 X APPROXIMATE REPEATS.
 FT REPEAT 146 322 1.
 FT REPEAT 341 540 2.
 FT REPEAT 341 540 3.
 FT DISULFID 5 14 BY SIMILARITY.
 FT DISULFID 27 40 BY SIMILARITY.
 FT DISULFID 39 51 BY SIMILARITY.
 FT DISULFID 73 118 BY SIMILARITY.
 FT DISULFID 117 126 BY SIMILARITY.
 FT DISULFID 149 195 BY SIMILARITY.
 FT DISULFID 194 203 BY SIMILARITY.
 FT DISULFID 216 232 BY SIMILARITY.
 FT DISULFID 231 242 BY SIMILARITY.
 FT DISULFID 266 311 BY SIMILARITY.
 FT DISULFID 310 321 BY SIMILARITY.
 FT DISULFID 344 390 BY SIMILARITY.
 FT DISULFID 389 396 BY SIMILARITY.
 FT DISULFID 409 425 BY SIMILARITY.
 FT DISULFID 424 435 BY SIMILARITY.
 FT DISULFID 485 531 BY SIMILARITY.
 FT DISULFID 530 539 BY SIMILARITY.
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 551 AA; 61455 MW; 2492404C8244E55 CRC64;

Query Match 11.8%; Score 365; DB 13; Length 551;

Best local similarity 23.0%; Pred. No. 9.3e-20;

Matches 134; Conservative 86; Mismatches 264; Indels 98; Gaps 17;

OY 53 CVADESANCDKSLHTLEFGDKICTYATLTRETYGEMADCCA-KOEPERNECFLOHKDNP 111
 DB 5 CCGENMAGCGLLHHRYTLFQDELCEGV--SIPSAACSSLANEBRADCLVSLR-GNLS 60
 OY 112 LPLRVREVDVWCTAFHD---NEETFLKTYELIARHHPYFVAPELLFFAKRYKAATE 167
 DB 61 IHSVPLAPASQLC--HDIRMKSHESF-ASLIMERGRHPRADSOVEELAEFRSKIGDA 116
 OY 168 CCOADRAKACLLPKLDELDEGKASSAKORLKASLOKFGERAFAKANAVALSORFPAE 227
 DB 117 CCDLADEKECITTRGREALHQEVSAYADADLCSSLAOLGAKFLGRVNVLFSGQRNAT 176
 OY 228 FAEYSKIVTDLTQVTECC-HGDLLEC-ADDRADLAKYICENODSIS--SKLKECEKPL 283
 DB 177 FDQISKLSHRFHSVAQOTCCGEGWSPGCFABQRIIHDMECHDMELSRVPMAMACCOISG 236
 OY 284 LEKSHCIAEVENDEKPADLPSELADPVESEKDYCKNYAKVDEKPLVEEPONLIKQNCLEFEOLGEY 343
 DB 237 SARAKCMETIRGRKRVLD--VALARFGHGVCCMNAAPQELLGRMLYEFGRRTDMSV 293
 OY 344 VLLRLAKTYETTLKCCAA--ADPHECYAKVDEKPLVEEPONLIKQNCLEFEOLGEY 401
 DB 294 GEAKKITEWDKQKDCAGNHSBQACLVSKKALISVIGEBQAKSHIKIEOLOKDHOE 353
 OY 402 KFONALVRYTKKYPQVSTPPLVEYSRNLGKVGSKCCNHPAKRPPCAEDLYSVLQDC 461
 DB 354 VFEKVVLDLFALEARTISLDKVEFAHRYTHAIFRCAH---QANCLLDENLHLSLIC 409
 OY 462 VLHETPVSDRYTKCCTSLVNRRCFSALEVDETYVEKNAEFTTHADICILSEKER 521

```

Db 410 SDLSYLAHDGYKCC-----RLAESEAVSCVHEHERAHEAR 447
QY 522 QIKK-QTALVELVKHKPKATKPOLKAVMDPFAA----- 553
Db 448 ATEEVENICKERVENQAKV--EAVEAVEPFAEBGAANCLFROLPGKYLORLLYKAH 505
QY 554 -----EVE---KCKADKDKETCEAECKKL 575
Db 506 QAPAGVDHSRIRLOVHHFEVYAKCCORAVDKSECSHEIKEM 547

RESULT 11
Q90WZ8 PRELIMINARY: PRT: 122 AA.
AC Q90WZ8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Serum albumin (Fragment).
OS Larus argentatus (Herring gull).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauia; Aves; Neognathae; Charadriiformes; Laridae; Larus.
OX NCBI_TaxID=35669;
RN [1]
RP SEQUENCE FROM N.A.
RA Lorenzen A., Casley W.L., Moon T.W.;
RT "Development of an RT-PCR Bioassay for Avian Vitellogenin mRNA.";
RL Toxicol. Appl. Pharmacol. 0:0-0(2001).
DR EMBL: AY045725; AAL01533.1;
DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transport_prot; 1.
FT NON_TER 1
FT NON_TER 122
SQ SEQUENCE 122 AA; 13805 MW; 93C644A0B120EF93 CRC64;

Query Match 10.7%; Score 331; DB 13; Length 122;
Best Local Similarity 47.1%; Pred. No. 5.9e-18;
Matches 57; Conservative 23; Mismatches 41; Indels 0; Gaps 0;

QY 453 YLSVVLNOLCVLHEKTPVSDRYTKCTESLVNRRPCFSALVDETYVPKEFNAETTFHA 511
Db 1 YLSYIQQDCRQETTPVNDVSHCCSDSYAVRRPCTFMGVDYKVPAPFPEMFSPDE 60
QY 512 DICTLSEKEROIKQTALVELVKHKPKATKPOLKAVMDPFAAVEKCKADDKETCFAE 571
Db 61 KICTAPPAEQLGQKLLINLKKRPQMTIEQIKITADGFTAMVDKCKQSDIEICFGE 120
QY 572 G 572
Db 121 G 121

RESULT 12
Q90WZ6 PRELIMINARY: PRT: 123 AA.
AC Q90WZ6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Serum albumin (Fragment).
OS Porphila guttata (Zebra finch) (Taeniopygia guttata).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauia; Aves; Neognathae; Passeriformes; Estrildidae;
OC Estrildinae; Taeniopygia.
OX NCBI_TaxID=59729;
RN [1]
RP SEQUENCE FROM N.A.
RA Lorenzen A., Casley W.L., Moon T.W.;
RT "Development of an RT-PCR Bioassay for Avian Vitellogenin mRNA.";
RL Toxicol. Appl. Pharmacol. 0:0-0(2001).
DR EMBL: AY045727; AAL01535.1;
DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transport_prot; 1.

```

```

FT NON_TER 1
FT NON_TER 123
SQ SEQUENCE 123 AA; 13824 MW; B18F83BC2E194F29 CRC64;

Query Match 9.3%; Score 290; DB 13; Length 123;
Best Local Similarity 43.0%; Pred. No. 8.2e-15;
Matches 52; Conservative 22; Mismatches 47; Indels 0; Gaps 0;

QY 453 YLSVVLNOLCVLHEKTPVSDRYTKCTESLVNRRPCFSALVDETYVPKEFNAETTFHA 511
Db 1 YLTVIIEQCKKQESTPVNDVSHCCNDYSKRCFTFMGDTYKVPAPFPTLDE 60
QY 512 DICTLSEKEROIKQTALVELVKHKPKATKPOLKAVMDPFAAVEKCKADDKETCFAE 571
Db 61 KCKAPPAERAGELKLLVNLVKRRPQMTIEQIKITBGFAMMECKCKPVEGCLGE 120
QY 572 G 572
Db 121 G 121

RESULT 13
Q63205 PRELIMINARY: PRT: 135 AA.
AC Q63205;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Messenger RNA for rat alpha-fetoprotein (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-81006964; Pubmed-6157690;
RA Innis M.A., Miller D.L.;
RT "Alpha-fetoprotein gene expression. Partial DNA sequence and COOH-terminal homology to albumin.";
RL J. Biol. Chem. 255:8994-8996(1980).
DR EMBL: V01236; CAA24546.1;
DR HSSP: P02768; IE7B.
DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transport_prot; 1.
DR SMART: SM00103; ALBUMIN; 1.
FT NON_TER 1
FT NON_TER 135
SQ SEQUENCE 135 AA; 15040 MW; 0F922DAFAF71EAC4 CRC64;

Query Match 8.5%; Score 264; DB 11; Length 135;
Best Local Similarity 41.0%; Pred. No. 8.9e-13;
Matches 50; Conservative 20; Mismatches 52; Indels 0; Gaps 0;

QY 450 EDYLSVVLNOLCVLHEKTPVSDRYTKCTESLVNRRPCFSALVDETYVPKEFNAETTFE 509
Db 1 EGLADIYIGHLCLRHANVNSGINHCCSSSNRRLCTISFLRRETYVPLPFSATNSS 60
QY 510 HADICTLSEKEROIKQTALVELVKHKPKATKPOLKAVMDPFAAVEKCKADDKETCF 569
Db 61 TRNLCOAGRAPQMTKQELLINLVKQKPMTEBOAAVADPSGLIECKCKDDQDQACFA 120
QY 570 EE 571
Db 121 KE 122

RESULT 14
Q9JMX8 PRELIMINARY: PRT: 1723 AA.
AC Q9JMX8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE Cag-Y.

```


GN CAG-Y.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CCUG 17874;
RA MEDLINE=97121442; PubMed=8962108;
RX Genist S., Lange C., Xiang Z., Crabtree J.E., Chiara P.,
RA Borodovsky M., Rappaport R., Covacci A.,
RT "cag, a pathogenicity island of Helicobacter pylori, encodes type I-
specific and disease-associated virulence factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:14648-14653(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CCUG 17874;
RA MEDLINE=20150112; PubMed=10684850;
RX Covacci A., Rappaport R.,
RT "Tyrosine-phosphorylated bacterial proteins: trojan horses for the
host cell.";
RL J. Exp. Med. 191:587-592(2000).
DR EMBL: AF282852; AAF80198.1;
SQ SEQUENCE 1723 AA; 196048 MW; 0FAC456B76622801 CRC64;

Query Match 6.1%; Score 188; DB 2; Length 1723;
Best Local Similarity 22.9%; Pred. No. 1.2e-05;
Matches 157; Conservative 88; Mismatches 254; Indels 188; Gaps 37;

OY 16 EENFALVILIAFAOYLQCPEDHVKLVNEYTEPAKTCVADSAENC-----DKSLHTLF 70
DB 351 EERIKCDLIDENLKLKSLNQKQVALDCLKNKKT--DEERNECKLINDPEIRKRF 407
OY 71 GDLCTVATLTRETYGEMADCC--AKQEPERNECFLOHDDNPNRLVREPVDVWCTAFH 128
DB 408 RKEL-----GLOKELOEYKDCIKNKTAEKNECKLGLSKE--AIERLQOALDCLKNKKT 461
OY 129 DNEETFLKLYEIAIRRHRYFAPELL--FFAKRYKAFTCCQADRAAC---LLPK-- 181
DB 462 DEERN-----ECKLNIPDOLQKELLADMSYKAVKDCVSRARNEKEKECEKLLTPPAK 514
OY 182 -----LDELDECKASSAKORLKAS-----LQKGEAFKAMAVARLSQFPP 224
DB 515 KLEDOVALDCLKN--AKTDERKKCLDLPKQLOSDILAKESLAKYDC--VSQAKT 567
OY 225 KAEEFAVSCLVY-----DLTKVHTEC-----CHGDILLECADRDADL 260
DB 568 EAEKKECEKLLTPPAKLLLEBAKESVAYLDVSOAKTEAKKECEKLLTPPAKLLLE 627
OY 261 AK-----YI-CENODSISKLKECE-----KPLLEKS--HCI--AEVEND-----EM 298
DB 628 AKKSVKAYLDVSOAKTEDEKKECEKLLTPPAKLLLEQALDCLKNKTAADKRCVADL 687
OY 299 PADLPS--LAAD-----FVESKDVCNVAEADVFLGMFLYEARRHPDYSVLLRLAKT 352
DB 668 PKDLQKVLAEESVRYLDVSRARNEAEERKE--CEKLLTPPAK-------LLEBAK 737
OY 353 YETTLERKCCAA--DPEHCYAKVDFEFLVEEPOMLIKONCELFQOL--GEYKQON 405
DB 738 SVKATKDCVSRARNEKEKECEKLLTPPAKLLLEBAKESVAYLDVSOAKTEDEKKECE 797
OY 406 ALLVYTRKVPQVSTPLVE-----VSRNIGVGSCKCK--PEAKRMPCAEDYLSVVLN 458
DB 798 KLLFPEAKKLLLEBAKESVAYLDVSOAKTEDEKKECEKLLTPPAKLL-----LE 847
OY 459 QL-----CVLHEKTPVSDRYTKCTESLVNRRPCFSALVEDYTVKPEFAEFTT-----FH 510
DB 848 QOALDCLKNKTAAD--KKRCVKDL-----PKDLQKVLAKKSVAY 887
OY 511 ADICTLSEKROIK--OTALVELVYHKPKATREOLKAVMDPAFVKKCCADKQETC-- 567
DB 888 LDCVSRARNEKEKECEKLLTPPAKLLLEBAKESLKAIKD-----CLSOARNEEERRACEK 943

OY 568 -----FAEKKVLVA-----SQA 581
DB 944 LITPEARKLLEBAKESVAYLDVSOA 970

RESULT 15
ID Q9ZLV0 PRELIMINARY; PRT; 1819 AA.
AC Q9ZLV0;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE CAG island protein.
GN JHP0476.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Dolg P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL: AE001481; AAD06047.1;
SQ SEQUENCE 1819 AA; 207505 MW; E76762B5A7FA371D CRC64;

Query Match 5.9%; Score 184; DB 16; Length 1819;
Best Local Similarity 21.8%; Pred. No. 2.7e-05;
Matches 150; Conservative 101; Mismatches 278; Indels 158; Gaps 34;

OY 16 EENFALVILIAFAOYLQCPEDHVKLVNEYTEPAKTCVADSAENC-----DKSLHTLF 70
DB 371 EERIKCDLIDENLKLKSLNQKQVALDCLKNKKT--DEERNECKLINDPEIRKRF 427
OY 71 GDLCTVATLTRETYGEMADCC--AKQEPERNECFLOHDDNPNRLVREPVDVWCTAFH 128
DB 428 RKEL-----ELOKELOEYKDCIKNKTAEKNECKLGLSKE--AIERLQOALDCLKNKKT 481
OY 129 DNEETFLKLYEIAIRRHRYFAPELL--FFAKRYKAFTCCQADRAAC---LLPK-- 183
DB 482 DEERN-----ECKLNIPDOLQKELLADMSYKAVKDCVSRARNEKEKECEKLLTPPAK 534
OY 184 ELRDE-----GKASSAKORLKAS-----LQKGEAFKAMAVARLSQFPAE 227
DB 535 KLENOALDCLKNKTADEKKECKLNPDKLOSDILAKESLAKYKDC--VSQAKTBAE 590
OY 228 FAEVSCLVY-----DLTKVHTEC-----CHGDILLECADRDADL 262
DB 591 KKECEKLLTPPAKLLLEBAKESVAYLDVSOAKTEAKKECEKLLTPPAKLLLEBAK 650
OY 263 -----YI-CENODSISKLKECE-----KPLLEKS--HCI--AEVEND-----MPAD 301
DB 651 SVRAYLDVSRARNEAEERKECEKLLTPPAKLLLEQALDCLKNKTADEKKECKLDPKD 710
OY 302 LPS--LAAD-----FVESKDVCNVAEADVFLGMFLYEARRHPDYSVLLRLAKT 355
DB 711 LQKVLAEESVRYLDVSRARNEAEERKE--CEKLLTPPAK-------LLEBAKSVK 760
OY 356 TLEKCCAA--DPEHCYAKVDFEFLVEEPOMLIKONCELFQOL--GEYKQON 408
DB 761 AYKDCVLRARNEKEKECEKLLTPPAKLLLEBAKESVAYLDVSRARNEAEERKECEKLL 820
OY 409 VRYTRKVPQVSTPLVE-----VSRNIGVGSCKCK--PEAKRMPCAEDYLSVVLN 460
DB 821 TPEARKLLEBAKESVAYKDCVSRAR--EKEKQCEKLLTPPAKLL--LEESKSVAY 878

THIS PAGE BLANK (uspto)